

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:08:57 ; Search time 2307.23 Seconds

(without alignments)
3998.551 Million cell updates/sec

Title: US-09-827-854-14

Perfect score: 1581
Sequence: 1 MKVLMALVTFELAGCQAKV.....VEKQAVGTSAVPSPDNH 317

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xip
-O=/cgn2.1/USPTO-spool/US09827854/runat_11032003_101610_27486/app-query.fasta.1.3576
-DB=genemb1 -QFMT=fastcap -SUFFIX=g2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=humand40.cdi -LIST=45
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854 -EGCN_1.1.1308 -etrunat_11032003_101610_27486 -NCPU=6 -ICPU=3
-NO_XLPTX -NO_MMAP -LARGEQUERY -NEG1SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hc:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rnd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1581	100.0	1156	6	BD004279 Apo E hum
2	1573	99.5	1110	6	E00359 cDNA coding
3	1573	99.5	1110	6	E00823 DNA sequence
4	1573	99.5	1147	6	AX302545 Sequence
5	1573	99.5	1156	6	BD004278 Apo E hum
6	1573	99.5	1156	9	H00396 Homo sapien
7	1573	99.5	1186	9	BC003557 Homo sapi
8	1566	99.1	1110	6	E08423 DNA coding
9	1565	99.0	1156	6	BD004277 Apo E hum
10	1564	98.9	1157	6	AX333278 Sequence
11	1564	98.9	1157	6	AX409597 Sequence
12	1564	98.9	1157	6	E15975 Sequence 1
13	1564	98.9	1157	9	H00406E Human apoli
14	1559	92.3	1178	9	MEAPOE Monkey mRNA
15	1387	87.7	5515	9	H00406E Human apoli
16	1384.5	87.6	5491	9	AF261279 Homo sapi
17	1384.5	87.6	41907	6	AX358722 Sequence
18	1384.5	87.6	41907	9	AF050154 Homo sapi
19	1384.5	87.6	107567	9	AC011481 Homo sapi
20	1358	85.9	5413	9	AF261280 Pan trogl
21	1275	80.6	4762	9	BABAPOE Baboon apoli
22	1274	80.6	208239	2	AC021988 Homo sapi
23	1266	80.1	1138	4	AF303830 Tupaia gl
24	1159	73.3	718	9	AF200499 Pan trogl
25	1158	73.2	1060	4	RABAPOLP Rabbit apoli
26	1157	73.2	1045	10	MUSAPOE Mus apoli
27	1157	73.2	1104	10	BC028816 Mus muscu
28	1155	73.1	718	9	AF200500S3 Gorilla g
29	1155	73.1	718	9	AF200500S3 Pongo pyg
30	1153.5	73.0	1154	4	BRAPOEHR B. taurus mr
31	1151.5	72.8	965	6	AX384545 Sequence
32	1151.5	72.8	1108	4	BRAPOLPE B. taurus mr
33	1151.5	72.8	5617	6	AX384541 Sequence
34	1151.5	72.8	6026	6	AX384539 Sequence
35	1151	72.8	718	9	AF200506S3 Hylobates
36	1151	72.8	1122	4	SSAPOE S. scrofa mr
37	1146	72.5	1126	6	AR164342 Sequence
38	1146	72.5	1126	6	AR205885 Sequence
39	1102	69.7	951	10	S76779 RAPOE-apol1
40	1080	68.3	959	10	MUSAPOE Mus musculu
41	1064.5	67.3	1069	10	RAIPAOE J00705 rat apoli
42	1037	65.6	228698	2	AC127479 Mus muscu
43	1037	65.6	237653	2	AC073760 Mus muscu
44	1011	63.9	4856	10	MUSAPE D00466 Mus musculu
45	1010	63.9	4267	6	AR164387 Sequence

RESULT 1

ALIGNMENTS


```

source          1..1110
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
BASE COUNT      198 a      353 c      416 g      143 t
ORIGIN
Alignment Scores:
Pred. No.:      1,95e-77      Length:      1110
Score:          1573.00      Matches:      316
Percent Similarity: 99.68%      Conservative: 0
Best Local Similarity: 99.68%      Mismatches: 1
Query Match:    99.49%      Indels:      0
DB:             6      Gaps:      0
US-09-827-854-14 (1-317) x E00359 (1-1110)
OY      1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB      15 ATGAAGGTTCTGTGGGCTGCTGGTGGTCACATTCTCGCAGATGCGAGGCCAAGGTG 74
OY      21 GtGlnAlaValGtThGluProGluProGluLeuArgGlnGlnThrGluThrPheGlnSer 40
DB      75 GAGCAAGCGGTGGAGACAGCGGAGCCGAGCTGGCCAGCAGACGAGCGAGAGC 134
OY      41 GlyLnaArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB      135 GGCCAGCGCTGGGAATGCGACATGGGTGCTTTGGGATTACTGCGCGGGTGCAGACA 194
OY      61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgala 80
DB      195 CTGTGTGAGCAGGTGAGAGAGAGTGTCTCAGCTCCAGAGTCACCCAGAACTGAGGGCG 254
OY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
DB      255 CTGATGAGCAGACCATGATGAGAGAGTTGAGGCTTACAAATCGAACTGGAGAGACACTG 314
OY      101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB      315 ACCCGGCGGCGGAGAGACGCGGCGACGCGCTGCCAAGAGACTGCGAGCGCGGCGAGCG 374
OY      121 ArgLeuGlyAlaAspMetClnAspValArgGlyArgLeuValGlnTyrArgGlyGluVal 140
DB      375 CGCGTGGGCGGACATGAGAGAGCTGTGCGCGCCGCTGTGCTACAGTACCGCGGCGAGGTG 434
OY      141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB      435 CAGGCCATGCTCGCCAGAGACACCGAGAGCTGCGGGGTGCGCTGCCCTCCACCTGCGCC 494
OY      161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB      495 AAGCTGCGTAAGCGGCTCTCGCGCATGCGCATGACCTGCACAAACGCGCTGCGCATGTAC 554
OY      181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
DB      555 CAGGCGCGGCGCGGAGGCGCGCCAGGCGCGCTCAGGCGCATCCGAGAGCGCTCGGGG 614
OY      201 ProLeuValGlnGlnGlyArgValArgValAlaThrValGlySerLeuAlaGlyGlnPro 220
DB      615 CCCCTGCTGGAGACAGGCGCGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 674
OY      221 LeuGlnGluArgAlaGlnAlaLeuTrpGlyGluArgGluArgAlaLeuMetGly 240
DB      675 CTACAGAGAGCGGCGCGAGCGCTGGGGGAGCGCGCTGCGCGCGCGATGGAGAGATGGGGC 734
OY      241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB      735 AACCCGAGCGCGGAGCGCGCTGGAGCAGGTGAAGGAGAGGTGGCGAGGTGGCGCCCAAG 794
OY      261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaLeuArgLeu 280
DB      795 CTGGAGAGAGCGCGCGCGAGATGCGCTGCGAGCGCGGAGGCTTCCAGGCGCGCTCAAG 854
OY      281 SerTrpPheGluProLeuValGlnLysPheGlnArgGlnTrpAlaGlyLeuValGluLys 300

```

```

DB      855 ACCTGGTTGAGACCCCTGTTGTAAGACATGACGCGCATGTGGCGGCTGTTGGAGAG 914
OY      301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB      915 GTGCAGGCTGCCGTGGGACACGAGCGCGCCCTGTGTCCAGCAGACATCAC 965
RESUT 3
E00823
LOCUS      E00823
DEFINITION DNA sequence coding for human apolipoprotein E and its signal
            peptide.
ACCESSION      E00823
VERSION      E00823.1 GI:2169084
KEYWORDS      JP 1986096997-A/1.
SOURCE      unidentified.
ORGANISM      unidentified.
REFERENCE      1 (bases 1 to 1110)
AUTHORS      Teranishi, Y., Matsui, Y., Ikeda, Y. and Kimura, M.
TITLE      PRODUCTION OF HUMAN APOLIPOPROTEIN E-LIKE PROTEIN
JOURNAL      Patent: JP 1986096997-A I 15-MAY-1986;
COMMENT      MITSUBISHI CHEM IND LTD
OS      Human {Homo sapiens}
PN      JP 1986096997-A/1
PD      15-MAY-1986
PF      16-OCT-1984 JP 1984216987
PI      TERANISHI YUTAKA, MATSUI YASUSHI, IKEDA YASUOKO, KIMURA MASAKO
PC      C12P21/00, A61K35/74, A61K37/04, C12N15/00, (C12P21/00, C12N1:19),
PE      (C12N15/00,
PP      C12R1:19);
PT      CC
PS      CC
STRANDEDNESS Double;
CC      topology: linear;
CC      hypothetical: No;
CC      anti-sense: No;
CC      *source: tissue= liver;
FH      Key      location/Qualifiers
FT      3'UTR      1..14
FT      sig-peptide      15..68
FT      /product='human apolipoprotein E signal' FT
FT      /peptide      69..968
FT      CDS
FT      969..1110.      /product='human apolipoprotein E' FT      3'UTR
FEATURES
source      1..1110
            location/Qualifiers
            /organism='unidentified'
            /db_xref='taxon:32644'
BASE COUNT      198 a      353 c      416 g      143 t
ORIGIN
Alignment Scores:
Pred. No.:      1,95e-77      Length:      1110
Score:          1573.00      Matches:      316
Percent Similarity: 99.68%      Conservative: 0
Best Local Similarity: 99.68%      Mismatches: 1
Query Match:    99.49%      Indels:      0
DB:             6      Gaps:      0
US-09-827-854-14 (1-317) x E00823 (1-1110)
OY      1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB      15 ATGAAGGTTCTGTGGGCTGCTGGTGGTCACATTCTCGCAGATGCGAGGCCAAGGTG 74
OY      21 GtGlnAlaValGtThGluProGluProGluLeuArgGlnGlnThrGluThrPheGlnSer 40
DB      75 GAGCAAGCGGTGGAGACAGCGGAGCCGAGCTGGCCAGCAGACGAGCGAGAGC 134
OY      41 GlyLnaArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB      135 GGCCAGCGCTGGGAATGCGACATGGGTGCTTTGGGATTACTGCGCGGGTGCAGACA 194

```


QY	61	LeuSerGIuGIuValGIuGIuGIuLeuSerSerGIuValThrGIuGIuLeuArgAla	80
Db	195	CTGCTCAGCAGGTCAGAGGAGGAGCTCTCAGCTCCAGGACACCGAGAACTGAGGGCG	254
QY	81	LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGIuLeu	100
Db	255	CTGATGAGCAGACCATGAAAGGAGATTGAAAGCCCTACAAATCGGAATCGGAGGAACAAC	314
QY	101	ThrProValAlaGIuGIuGIuThrArgAlaArgLeuSerLysGIuLeuGIuAlaGIuAla	120
Db	315	ACCCCGGTGGCGAGAGAGACGGCGGACGGCTGTCCAAAGAGCTGCAGGGCGCCAGAGCC	374
QY	121	ArgLeuGIuAlaAspMetGIuAspValArgGIuArgLeuValGlnTyrArgGIuVal	140
Db	375	CGGCTGGCGCGGACATGAGAGAGCTGTGCGCGCCCTGTGGCAGTACCGCGGAGAGTGC	434
QY	141	GIuAlaMetLeuGIuGIuInsThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg	160
Db	435	CAGGCCATGTCTGGCCAGACACCGAGAGAGCTGGGGGTGGCCCTCCGCTCCACCTGGCG	494
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGIuLysArgLeuAlaValTyr	180
Db	495	AAGCTGGCTTACCGGCTCTCCCGGATGCCGATGACTCTGCAGAAAGCCCTGGCAGGTAC	554
QY	181	GIuAlaGIuAlaArgGIuGIuGIuArgGIuArgGIuLeuSerAlaIleArgGIuArgLeuGI	200
Db	555	CAGGCCGGGGCCCGGAGGGGGCGGAGCGGCGCTCAGCGCCATCCGCGAGCGCTGGGG	614
QY	201	ProLeuValGIuGIuGIuArgValArgValAlaThrValGIuSerLeuAlaGIuGIuPro	220
Db	615	CCCTGTGTGAACAGGGCCCGCTGGCGGCGCCACTGTGGCTCCCTGGCGCGGCGACCGC	674
QY	221	LeuGIuGIuArgAlaGIuGIuArgLysGIuArgLeuArgAlaArgMetGIuGIuMetGI	240
Db	675	CTACAGAGAGGGGCCAGGCGCTGGGGGAGCGGCTGGCGCGGAGTGGAGAGATGGCG	734
QY	241	SerArgThrArgAspArgLeuAspGIuValLysGIuGIuValAlaGIuAlaArgAlaLys	260
Db	735	AGCGGAGCCCGGCGACCGCTCGAGAGAGTGAAGAGAGAGTGGCGGAGGTGGCGCCAA	794
QY	261	LeuGIuGIuGIuAlaGIuGIuGIuIleArgLeuGIuAlaGIuAlaPheGIuAlaArgLeuLys	280
Db	795	CTGAGAGAGAGAGCCAGCAGATACGCTTCAGGCCAGGCTTCAGAGGCCCTCCAAAG	854
QY	281	SerTyrPheGIuProLeuValGIuAspMetGIuArgGIuTyrPheGIuAlaGIuLys	300
Db	855	AGCTGGTTCAGGCCCTCGGTGGAAGACATCAGAGGCCAGTGGGCCGGGCTGTGGAGAG	914
QY	301	ValGIuAlaIleValGIuThrSerAlaIleProValProSerAspAsnHis	317
Db	915	GTCAGAGCTCGGTGGGACACGCGGCCCTGTGGCCAGCGACATATCAG	965
RESULT 4			
AX302545		1147 bp	DNA
LOCUS	AX302545		linear
DEFINITION	Sequence 63 from Patent WO01.75177.		PAT 30-NOV-2001
ACCESSION	AX302545		
VERSION	AX302545.1		GI:17383082
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Morin, P. J., Sherman-Baust, C. A., Pizer, E. S. and Hough, C. D.		
JOURNAL	Tumor markers in ovarian cancer		
FEATURES	Patent: WO 0175177-A 63 11-OCT-2001;		
source	THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)		
	Location/Qualifiers		
	1..1147		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		

BASE COUNT	210 a	365 c	425 g	147 t
ORIGIN				
Alignment Scores:				
Pred. No.:	2,02e-77	Length:	1147	
Score:	1573.00	Matches:	316	
Percent Similarity:	99.68%	Conservative:	0	
Best Local Similarity:	99.68%	Mismatches:	1	
Query Match:	99.49%	Indels:	0	
DB:	6	Gaps:	0	

US-09-827-854-14 (1-317) x AX302545 (1-1147)

QY	1	MetLysValLeuTyrPheAlaIleLeuValThrPheLeuAlaGIuCysGIuAlaLysVal	20
Db	46	ATGAGGTTCTGTGGGCTGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	105
QY	21	GIuGIuAlaValGIuThrGIuProGIuProGIuLeuArgGIuGIuGIuThrGIuTyrGIuIns	40
Db	106	GAGCAACCGGTGAGACAGAGCCGAGAGCCGAGACTCGCGCAGAGACCGAGTGGCAGAGC	165
QY	41	GIuGIuArgTyrPheGIuLeuAlaLeuGIuArgPheTyrPheTyrPheTyrPheTyrPheTyr	60
Db	166	GCCAGCGCTGGGAGACTGCGACTGCGCTTGGGATTTACCTGCGTGGTGGTGGTGGTGG	225
QY	61	LeuSerGIuGIuGIuGIuGIuGIuLeuLeuSerSerGIuValThrGIuGIuLeuArgAla	80
Db	226	CTGTCTGAGCAGGTGACAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	285
QY	81	LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGIuLeu	100
Db	286	CTGATGAGCAGACCATGAAAGGAGTGAAGGCCCTACAAATCGGAACCTGGAGGAACAAC	345
QY	101	ThrProValAlaGIuGIuGIuThrArgAlaArgLeuSerLysGIuLeuGIuAlaGIuAla	120
Db	346	ACCCCGGTGGGAGAGAGCGGCGGAGCTGTCCAGAGAGCTGCAGGCGGCGGCGAGGCC	405
QY	121	ArgLeuGIuAlaAspMetGIuAspValArgGIuArgLeuValGIuTyrArgGIuGIuVal	140
Db	406	CGGCTGGCGCGGACATGAGAGAGCTGTGGCGCGCTGTGGTGCAGTACCGCGGCGAGGT	465
QY	141	GIuAlaMetLeuGIuGIuInsThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg	160
Db	466	CAGGCCATGTCTGGCCAGACACCGAGAGCTTCGGGTGGGCTCCCTCCACCTGGCGCG	525
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGIuLysArgLeuAlaValTyr	180
Db	526	AGCTGGCTTACCGGCTCTCCCGGATGCCGATGACTCTGCAGAAAGCCCTGGCAGGTAC	585
QY	181	GIuAlaGIuAlaArgGIuGIuGIuArgGIuArgGIuLeuSerAlaIleArgGIuArgLeuGI	200
Db	586	CAGGCCGGGGCCCGGAGGGCGCGGAGCGGCGCTCAGCGCCATCCGCGAGCGGCGGCG	645
QY	201	ProLeuValGIuGIuGIuArgValArgValAlaThrValGIuSerLeuAlaGIuGIuPro	220
Db	646	CCCTGTGTGAACAGGCGCGCTGGCGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	705
QY	221	LeuGIuGIuArgAlaGIuGIuArgLysGIuArgLeuArgAlaArgMetGIuGIuMetGI	240
Db	706	CTACAGAGAGCGGCGCAGGCGCTGGGGGAGGGCTGGCGCGGAGTGGAGAGAGAGGCG	765
QY	241	SerArgThrArgAspArgLeuAspGIuValLysGIuGIuValAlaGIuAlaValArgAlaLys	260
Db	766	AGCGGAGCCCGGACCGGCTCGAGAGAGTGAAGAGACAGTGGCGGAGGTGGCGGCGCA	825
QY	261	LeuGIuGIuGIuAlaGIuGIuGIuIleArgLeuGIuAlaGIuAlaPheGIuAlaArgLeuLys	280
Db	826	CTGAGAGAGAGCCAGCAGATACCTTCAGAGCGGAGGCTTCAGGCGCGGCTCCAAAG	885
QY	281	SerTyrPheGIuProLeuValGIuAspMetGIuArgGIuTyrPheGIuAlaGIuLys	300
Db	886	AGCTGGTTCAGGCCCTCGGTGGAAGACATGAGCGCCAGTGGCGGCGGCTGTGGAGAG	945

Qy 301 ValGlnAlaIaValaGlyThrSerAlaIaProValProSerAspAsnHis 317
 |||||
 Db 946 GTGCAGGCTCGCGTGGGACACAGCGCCCTGTGTGCCAGGACAAATCAC 996
 RESULT 5
 BD004278
 LOCUS Apo E humanized mammal.
 DEFINITION BD004278
 ACCESSION BD004278.1 GI:18632239
 VERSION JP 2001017028-A/2.
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 1156)
 AUTHORS Fujita,S., Hamanaka,H.,
 TITLE Apo E humanized mammal
 JOURNAL Patent: JP 2001017028-A 2 23-JAN-2001;
 MITSUBISHI CHEMICAL CORP
 OS Homo sapiens (human)
 PN JP 2001017028-A/2
 PD 23-JAN-2001
 PE 28-APR-2000 JP 2000128919
 PR
 PI SHINOBU FUJITA,HIROKI HAMANAKA,YUKO FUKUI,MINESUKE YOKOYAMA PC
 A01667/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10, PC
 C12N15/09//C07K14/775,
 PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,(C12N5/00,C12R1:91) CC
 FEATURES
 source Location/Qualifiers
 FT Key Location/Qualifiers
 CDS (61)..(1011).
 1..1156
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 208 a 368 c 432 g 148 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.03e-77 Length: 1156
 Score: 1573.00 Matches: 316
 Percent Similarity: 99.68% Conservative: 0
 Best Local Similarity: 99.68% Mismatches: 1
 Query Match: 99.49% Indels: 0
 DB: Gaps: 0
 US-09-827-854-14 (1-317) x BD004278 (1-1156)
 Qy 1 MetLysValLeuTRPAlaIaLeuValThrPheLeuAlaGlyCysGlnAlaIaLysVal 20
 |||||
 Db 61 ATGAAGGTTCTGTGGCTGGCGTGTCTGGTCACATTCTCTGGCGAGATGCCAGGCTG 120
 Qy 21 GlnGlnAlaValaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpLysIns 40
 |||||
 Db 121 GAGCAAGGCGGTGAGACAGAGCGGAGCGCGAGCTGCCGACACACAGAGTGGAGAGC 180
 Qy 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 |||||
 Db 181 GCCCGAGCTGGAGCACTGGCTGGCTGCTTTGGATTACCTCCGCTGGGTCAGACACA 240
 Qy 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||||
 Db 241 CTGTCTGAGCAGGTCCAGAGAGCTGTCAAGTCCCAAGGTCACCCAGGAACCTGAGGCG 300
 Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGluGlnLeu 100
 |||||
 Db 301 CTGATGGACAGACCATGAGAGGTGAAGGCTTCAAAATCGAGACTGGAGAACAACTG 360
 Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaLysAla 120
 |||||
 Db 361 ACCCGGTGTGGGAGAGACAGCGGGCAGCGCTGTCCAAGGAGAGCTCAGAGCGGCGCAGCC 420

Qy 121 ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTrpArgGlyGluVal 140
 |||||
 Db 421 CGGCTGGGCGCGGACATATGAGAGACTGTGGCGCCCTGTGTGCATATACCGCGGAGAGTG 480
 Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||
 Db 481 CAGGCGATGCTCGGCGACAGACACCCAGAGAGCTGGGCTGTCCCTCCACCTGGCGC 540
 Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTrp 180
 |||||
 Db 541 AAGCTGCTTAAGGGCTCTCTCCGATGCTCCGATGCTGCAAGACCGCTGGCACTGTAC 600
 Qy 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgGly 200
 |||||
 Db 601 CAGCGCGGGCGCCGAGAGCGCGCGAGCGGCTTCACCGCATTCGGAGAGCGCTGGGG 660
 Qy 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyPro 220
 |||||
 Db 661 CCCCTGTGTGAACAGAGGCGGCTGGCGCGCCGCTGTGGCTCTCCCTGGCCGACGCCG 720
 Qy 221 LeuGlnGluArgAlaGlnAlaTrpGlyLysArgLeuArgAlaArgMetGluGlnMetGly 240
 |||||
 Db 721 CTACAGAGAGCGGCGCAGGCTGGGGCGAGCGGCTGCCGCGGATGAGAGATGGGC 780
 Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 |||||
 Db 781 AGCGCGAGCCCGGAGACCGCTGGAGCAGGTGAAGAGCAGTGGCGAGAGTGGCGCCAA 840
 Qy 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 |||||
 Db 841 CTGCAGAGAGAGCGCCACAGATACGCTGCAGCGCCGAGCGCTTCCAGCGCCGCTCAAG 900
 Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 |||||
 Db 901 AGCTGTGTGAGCCCTGTGTGGAACATGACAGCCGCTGTGGCTGTGTGAGAG 960
 Qy 301 ValGlnAlaIaValaGlyThrSerAlaIaProValProSerAspAsnHis 317
 |||||
 Db 961 GTGCAGGCTCGCGTGGGACACAGCGCCCTGTGTGCCAGGACAAATCAC 1011
 RESULT 6
 HUMAPOE3 1156 bp mRNA linear PRI 24-NOV-2000
 LOCUS HUMAPOE3
 DEFINITION Homo sapiens preapolioprotein E (APOE) mRNA, complete cds.
 ACCESSION K00396
 VERSION K00396.1 GI:178850
 KEYWORDS apolioprotein; apolioprotein E; lipoprotein; polymorphism; very low density lipoprotein.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 355 to 1156)
 AUTHORS Breslow,J.L., McPherson,J., Nussbaum,A.L., Williams,H.W.,
 TITLE Identification and DNA sequence of a human apolioprotein E cDNA clone
 JOURNAL J. Biol. Chem. 257 (24), 14639-14641 (1982)
 MEDLINE 83082756
 PUBMED 6897404
 2 (bases 250 to 777)
 REFERENCE
 AUTHORS Wallis,S.C., Rogne,S., Gyll,L., Markham,A., Edge,M., Woods,D.,
 TITLE The isolation of cDNA clones for human apolioprotein E and the detection of apoE RNA in hepatic and extra-hepatic tissues
 JOURNAL EMBO J. 2 (12), 2369-2373 (1983)
 MEDLINE 84131952
 PUBMED 6199196
 REFERENCE
 3 (bases 1 to 1156)
 AUTHORS Zannis,V.I., McPherson,J., Goldberger,G., Karathanasis,S.K. and Breslow,J.L.
 TITLE Synthesis, intracellular processing, and signal peptide of human apolioprotein E

JOURNAL J. Biol. Chem. 259 (9), 5495-5499 (1984)
 MEDLINE 84185684
 PUBMED 6325438
 REFERENCE 4 (bases 88 to 1156)
 AUTHORS McLean,J.W., Elshourbagy,N.A., Chang,D.J., Mahley,R.W. and Taylor,J.M.
 TITLE Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing of a new variant
 JOURNAL J. Biol. Chem. 259 (10), 6498-6504 (1984)
 MEDLINE 84212473
 PUBMED 6327682
 REFERENCE 5 (bases 577 to 624)
 AUTHORS Gill,L.L., Peoples,O.P., Pearson,D.H., Robertson,F.W., Humphries,S.E., Cumming,A.M. and Hardman,N.
 TITLE Isolation and characterisation of a variant allele of the gene for human apolipoprotein E
 JOURNAL Biochem. Biophys. Res. Commun. 130 (3), 1261-1266 (1985)
 MEDLINE 85279526
 PUBMED 2992507
 REFERENCE 6 (sites)
 AUTHORS Rall,S.C. Jr., Newhouse,Y.M., Clarke,H.R., Weisgraber,K.H., McCarthy,B.J., Mahley,R.W. and Bersot,T.P.
 TITLE Type III hyperlipoproteinemia associated with apolipoprotein E phenotype E3/3. Structure and genetics of an apolipoprotein E3 variant
 JOURNAL J. Clin. Invest. 83 (4), 1095-1101 (1989)
 MEDLINE 89198059
 PUBMED 2539388
 COMMENT [1] corrected in [J. Biol. Chem. 258, 11422-11422 (1983)]. [J. Biol. Chem. 258, 11422-11422 (1983)] correction of [1]. [4] epsilon-3 and variant.
 [5] epsilon-2 allele.
 [6] sites; mutations resulting in type III hyperlipoproteinemia. Apo E is a component of normal human very low density lipoprotein. There are six human apo E phenotypes known to result from a single structural gene, three of the common alleles being epsilon-4, epsilon-3 and epsilon-2. This sequence appears to be of the epsilon-3 allele. [1] argues that the apo E polymorphism involves mutations in the structural coding region; for example the epsilon-2 phenotype which is characterized by hyperlipoproteinemia is thought to result from a c to t change (arg to cys) at base 586 below [3],[5]. The sequence shown is 578 homologous with human apo A-I and 818 homologous with rat apo E. For the epsilon-4 sequence, see the separate entry.
 [J. Biol. Chem. 258, 11422-11422 (1983)] is too extensively revised relative to [1] to record the
 revisions in the FEATURES table. The sequence below is that of [J. Biol. Chem. 258, 11422-11422 (1983)]
 and [3].
 Apo E is located on chromosome 19 --Jackson, Bruns and Breslow , PNAS USA 81, 2945-2949 (1984) -- and is believed to be linked to the Apo C-II gene (see separate entry).
 The two mutations causing type III hyperlipoproteinemia (apolipoprotein E phenotype E3/3) produces substitutions of Arg for Cys at amino acid 112 and Cys for Arg at amino acid position 142. Draft entry and printed copy of sequence for [1] kindly provided by L.L. Gill, University of Aberdeen.
 Complete source information:
 Human liver [1] [J. Biol. Chem. 258, 11422-11422 (1983)], [2], [3], [4] and blood [5], cDNA to mRNA.
 Location/Qualifiers
 1. 1156
 /Organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="19q13.2"
 /tissue-type="liver and blood"
 1. 1156
 /gene="APOE"
 61. 1014
 /gene="APOE"
 /codon_start=1
 /product="preapolipoprotein E"
 /protein_id="AAB59546.1"

/db_xref="GI:178851"
 /db_xref="CDB:G00-119-691"
 /translation="MKVLMAALLVTFIACGAKVEQAVETEPPELRQOTEMSGORM
 ELAIGFPMDFLVNVOITLSQVOEELISSQVQELRLAMDETKEIKAYSELEDLTP
 VAEEFRARISKELOAAQARLGDMDVEYCRGLVOYRGVQAMLGOSTEELRVLASILR
 KLRRLRLDADLDQRLAVYQAGARGARGLSAIREIRGLPLVEGRVRAVGSILAG
 QPLQERNAQMGRLRLARMEMGSRRTDRIDEVKEVQAEVRAKLEQAOQIRLOAEAFQ
 ARKSWFEPLVEDMQRWAGLVEKVOAAVGTSAAPVPSDNH"
 61. 114
 /gene="APOE"
 /note="apolipoprotein E"
 115. 1011
 /gene="APOE"
 /product="apolipoprotein E mature peptide"
 267
 /gene="APOE"
 /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3], [4]"
 /citation=[2]
 /replace="c"
 279
 /gene="APOE"
 /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3]"
 /citation=[4]
 /replace="a"
 288
 /gene="APOE"
 /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3]"
 /citation=[4]
 /replace="a"
 369
 /gene="APOE"
 /note="in variant"
 /citation=[4]
 /replace="a"
 409
 /gene="APOE"
 /note="in variant"
 /citation=[4]
 /replace="a"
 448
 /gene="APOE"
 /note="in type III hyperlipoproteinemia"
 /replace="c"
 538
 /gene="APOE"
 /note="in type III hyperlipoproteinemia"
 /replace="t"
 568
 /gene="APOE"
 /note="in variant"
 /citation=[4]
 /replace="c"
 586
 /gene="APOE"
 /note="in epsilon-2 allele"
 /citation=[3]
 /citation=[5]
 /replace="t"
 702
 /gene="APOE"
 /note="c in [J. Biol. Chem. 258, 11422-11422 (1983)], [3], [4]"
 /citation=[2]
 /replace="g"
 783
 /gene="APOE"
 /note="in variant"
 /citation=[4]
 /replace="t"
 858
 /gene="APOE"
 /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3]"

Alignment Scores:

Pred. No.:	2,096-77	Length:	1186
Score:	1573.00	Matches:	316
Percent Similarity:	99.68	Conservative:	0
Best local Similarity:	99.68	Mismatches:	1
Query Match:	99.498	Indels:	0
DB:	9	Gaps:	0

US-09-827-854-14 (1-317) x BC003557 (1-1186)

QY	1	MeLysValLLeuTPPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	46	ATGAGGTTCTGTGTGGGCTGCTCTGGTCAATTCTCTGGAGGATGCCAGGCCAAGGTG	105
QY	21	GIuGlnAlaValAGluThrGluProGluProGluLeuArgGlnInpThrGluTrpGlnSer	40
Db	106	GAGCAAGCGGTGGAGACAGACCGCGAGCTGGCGCACGACGACCGAGTGGCANAGC	165
QY	41	GlyGlnArgTrpGluLeuAlaLeuGlyLysArgPheTrpAspTyrLeuArgTrpValGlnThr	60
Db	166	GGCCAGCGCTGGGAACATGGCACTGGGTGCGCTTTTGGGATTACCTGGCTGGGTCCAAACA	225
QY	61	LeuSerGIuGlnValGlnGluLeuLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80
Db	226	CTGTCTGAGCAGGTGCAGAGAGAGACTGTCAAGCTTCCAGTACGCCAGGAACGTAGGGCG	285
QY	81	LeuMetAspGIuThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu	100
Db	286	CTGATGGACGAGACCATGAAGAGATTGAAGGCCCTCAAAATCGGAACGTGAGGAACACTG	345
QY	101	ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	346	ACCCCGGTGGGAGAGAGACCCGGGCAAGGCTGTCCAAAGAGCTGACAGCGCGGCACAGCC	405
QY	121	ArgLeuGlnAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyGluVal	140
Db	406	CGCGTGGGGCGGACATGAGAGACGTGTGGCGCCCTGTGTGCATGACCGGGGAGAGTG	465
QY	141	GlnAlaMetLeuGlnGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	466	CAGGCCATGCTCGGCGCACAGAGACCGAGAGACTGGGGGTGGCGCTCGCTCCACACTGGCG	525
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	526	AAGCTGGGTAAAGCGGCTCTCCGGAATCCGATACCTGCAGAAAGCCCTCGCACAGTGTAC	585
QY	181	GlnAlaGlyAlaArgGluGluGluArgGlyLeuSerAlaLeuArgGluArgLeuGly	200
Db	586	CAGCGCCGGGGCCCGAGAGGGGGCGGACGGCGCTCAAGCCCATCCGCGAGCGCTGTGGG	645
QY	201	ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro	220
Db	646	CCCCTGGTGAACAGAGGGCCGCTGCGGGCGCCACTGTGGGTCTCCGTGGCGGCCAGCGC	705
QY	221	LeuGlnGluArgAlaGlnAlaTyrGlyLysArgLeuArgAlaArgMetGluGluMetGly	240
Db	706	CTACAGAGAGGGGCCCAAGCGCTGGGGGAGACGGCTGCGCGCGGATGAGAGAGATGGCG	765
QY	241	SerTrpThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys	260
Db	766	AGCCGGACCCCGACCCCTGTGAGAGAGTGAAGAGACAGGTGGCGGAGGTGCGGCCAGAG	825
QY	261	LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluLysPheGlnAlaArgLeuLys	280
Db	826	CTGGAGGAGCAGGCCCAAGCATACGCTCTCAGGCCAGAGCCCTTCCAGCGCCGCTCCAG	885
QY	281	SerTrpPheGluProLeuValGluAspMetGlnArgGlnIleTrpAlaGlyLeuValGluLys	300
Db	886	AGCTGATTCCAGCCCTGTGGTGAAGACATCAGCGCCACATGGGCGGTGTGGTGAAG	945
QY	301	ValGlnAlaAlaValAlaGlyTyrSerAlaAlaProValProSerAspHis	317
Db	946	GTCGAGAGCTCCGTGGGACACAGCGCCGCTGTGGCCAGCACACATATCA	996

<hr/>					
RESULT 8	E08423	1110 bp	RNA	linear	PAT 29-SEP-1997
LOCUS	E08423	DNA coding human apolipoprotein E3.			
DEFINITION	E08423				
ACCESSION	E08423.1	GI:2176540			
VERSION	JP 1994315392-A/L.				
KEYWORDS	Homo sapiens.				
SOURCE	Homo sapiens.				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	1 (bases 1 to 1110)				
TITLE	Morimoto, H. and Teranishi, Y.				
JOURNAL	METHOD FOR PRODUCING APOLOPROTEIN				
	Patent: JP 1994315382-A 1 15-NOV-1994;				
	MITSUBISHI KASEI CORP				
COMMENT	OS Homo sapiens (human)				
	PN JP 1994315392-A/1				
	PD 15-NOV-1994				
	PE 11-JUN-1985 JP 1994015433				
	PI MORIMOTO HIROMORI, TERANISHI YUTAKA				
	PC C12P21/02,C07K13/00,C12N5/10,C12N15/12,C12N15/18,(C12P21/02,				
	PC C12R1:91),				
	PC (C12N5/10,C12R1:91);				
	CC strandedness: Double;				
	CC topology: Linear;				
	FH Key Location/Qualifiers				
	FH source 1..1110				
	FT sig_peptide /organism='Homo sapiens'				
	FT mat_peptide 15..68				
	FT /product='human apolipoprotein E3'.				
FEATURES	Location/Qualifiers				
source	1..1110				
	/organism='Homo sapiens'				
	/db_xref='taxon:9606'				
BASE COUNT	198 a 355 c 414 g 143 t				
ORIGIN					
<hr/>					
Alignment Scores:					
Pred. No.:	4.67e-77	Length:	1110		
Score:	1566.00	Matches:	315		
Percent Similarity:	99.37%	Conservative:	0		
Best Local Similarity:	99.37%	Mismatches:	2		
Query Match:	99.05%	Indels:	0		
DB:	6	Gaps:	0		
<hr/>					
US-09-827-854-14 (1-317) x E08423 (1-1110)					
OY	1	MettysValleuTTPAlalaLeuValnThrPhetaleuaglycysGlnAlaLysVal	20		
Db	15	ATGAAGGTTCGTGGCGCTGCTGTCATTCCGCAGAGATGCCAGCAAGTG	74		
OY	21	GlunAlaValaLgUthrGluPrGluProLuleuarGngInThrGluTrpGlnSer	40		
Db	75	GAGCAAGCGGTGGACAGACCAGCCAGCCAGCTGCCCGACAGACCGATGGCAGAC	134		
OY	41	GlyGlnArGTtPGluleuAlaLeuGLyARphETrpASpyrLileuArGTtPvalGlnThr	60		
Db	135	GGCCAGCGCTGGAAACTGGCACTGGCTGCTTTGGATWACCTGCCCTGGGTCACACA	194		
OY	61	LeusertGluInValGlnGlnGluLuleuLenserGlnValIThrngInGluLeuArgLa	80		
Db	195	CTGTCTGAGCAGGTGGACAGAGACTGCTCAGCTCCAGGTACCCAGAACCTAGGGCG	254		
OY	81	LeuMetAPGluTrnMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu	100		
Db	255	CTGATGGACGACACCATTAAGAGAGTTGAAGCCCTTCAAAATGGAACTGGAGAACCACTG	314		
OY	101	ThrProVALaLagLugLuTrhArGalaaRglseuSerLysGluLeuGlnAlaLagLinaLa	120		

Db	315	ACCCGGTGGGGGAGAGACAGCGGGCCAGCGCTGTCTCAAGAGACTCTCAGCGGGCCAGCGCC	374
Oy	121	ArgLeuGlyAlaAspMetGlnAspValArgGlyArgLeuValGlnTyrArgGlyGluVal	140
Db	375	CGGCTGGGGCGGAGCATGGAGGACTGTGGGGCCGCCCTGGTGCATTTACCGGGCGAGGTG	434
Oy	141	GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	435	CAGGCCATGCTCGCGCAAGACACGAGAGAGCTGGGGTGGCCCTGCTCCACACTCGGC	494
Oy	161	LysLeuArgGlyArgLeuLeuArgAspAlaAspSerLeuGlnLysArgLeuAlaValLyr	180
Db	495	AAGCTGCTTAAGCGGCTCTCTCCGCAATCCCATGACCTGCAGAAAGCGCTGGCACTTAC	554
Oy	181	GlnAlaGlyAlaArgGluGlnValAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200
Db	555	CAGCGCGGGGGCCCGGAGAGGCCCCAGCGCGGGCTCTCACGCCCATCCGGAGCGCTGGGG	614
Oy	201	ProLeuValGluGlnGlyArgValArgAlaAlaIleThrValGlySerLeuAlaGlyGlnPro	220
Db	615	CCCCGTGTGAACAGAGCGCGCTGCGGGCCGCCACTGTGGCTCCCTGGCGGCGACGG	674
Oy	221	LeuGlnGluArgAlaGlnAlaIleProGlyLysLysArgLeuAlaArgMetGluMetGly	240
Db	675	CTACAGAGAGCGGGCCCAAGGCTCTGGGGCGAGCGGCTGGCGCGGATGAGAGATGGGC	734
Oy	241	SerArgTyrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys	260
Db	735	AGCCGGACCCCGAGACCGCTGGACAGAGGTGAAGAGACAGGTGGCCGAGGTGCGGCCAAG	794
Oy	261	LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaIleGlnAlaArgLeuLys	280
Db	795	CTGGAGGAGCAGGCGCCACAGATACGCTGCAGCGCGGACCTTCAGCGCCGCTCAAG	854
Oy	281	SerTyrPheGluProLeuValGluAspMetGlnArgGlnTyrPheAlaGlyLeuValGluLys	300
Db	855	AGCTGGTTCGAGCCCTGTGTGGAAACATGTGCAGCCCACTGTGGCGGCTGTGTGAGAG	914
Oy	301	ValGlnAlaAlaValAlaGlyThrSerAlaIleAlaProValProSerAspAsnHis	317
Db	915	GTGCAGGCTGCCGTGGGCACACGCGCCGCCCTGTGTGCCACAGCAGCAATCAC	965
RESULT 9	BD004277		
LOCUS	BD004277	1156 bp	DNA
DEFINITION	Apo E humanized mammal.		Linear
ACCESSION	BD004277		
KEYWORDS	BD004277.1 GI:18632238		
SOURCE	JP 2001017028-A/1.		
ORGANISM	Homo sapiens.		
REFERENCE			
AUTHORS			
TITLE	Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.		
JOURNAL	Patent: JP 2001017028-A 1 23-JAN-2001;		
COMMENT			
	OS Homo sapiens (human)		
	PN JP 2001017028-A/1		
	PD 23-JAN-2001		
	PF 28-APR-2000 JP 2000128919		
	PR		
	PI SHINOBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA PC		
	CI1N167/027,A61K45/00,A61P25/28,A61P43/00,CI2N5/10, PC		
	PC (CI2N5/10, CI2R1:91), CI2N5/00, CI2N15/00, (CI2N5/00, CI2R1:91) CC		
FEATURES			
SOURCE	FT Key Location/Qualifiers		
	CDS (61)..(1011).		
	1..1156		
	/organism="Homo sapiens"		

QY 301 ValGlnAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCGTGGGACACGCGCCCTGTGCTGCACGACATCTAC 1011
RESULT 10
AX333278
LOCUS AX333278 1157 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 3787 from Patent WO0194629.
ACCESSION AX333278
VERSION AX333278.1 GI:18123912
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,
Horrikan, S., Soppet, D. R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 3787 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source location/Qualifiers
1..1157
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN
Alignment Scores:
Pred. No.: 6.26e-77 Length: 1157
Score: 1564.00 Matches: 314
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 99.05% Mismatches: 3
Query Match: 98.92% Indels: 0
Gaps: 0
DB: 6
US-09-827-854-14 (1-317) x AX333278 (1-1157)
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAGGTTCTGTGGCTGTGGCTGTGCATCTCTGCGAGATGCCAGGCAAGGTG 121
QY 21 GlnGlnAlaValAlaGlyThrGlnProGlnProGlnLeuArgGlnGlnThrGlnThrPgiIns 40
Db 122 GACCAAGCGGTGAGAGACAGAGCGGAGCCGAGCTGCGCAGAGACCGAGTGCAGAGC 181
QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 182 GGGCAGCGCTGGAGACTGGCTGCTGCTTTGGATTACCTGCGCTGCGCAGACA 241
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db 242 CTGTCTGAGCAGGTGCAAGAGAGCTGCTCAGCTCCCAAGTCCCAAGAGACTGAGGCG 301
QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
Db 302 CTGATGAGAGACATGAAGAGCTTGAAGGCTTACAAATCGAATCGAGAGAACACTG 361
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
Db 362 ACCCGGTAGCGAGAGAGAGCGGCGGCTGTCCAAAGAGCTGCAAGAGCTGCAAGCGG 421
QY 121 ArgLeuGlnAlaAspMetGlnAspValArgGlyArgLeuValGlnThrArgGlyGlnVal 140
Db 422 CGGCTGGGCGGAGACATGAAGAGCTGTGCGGCGGCTGTGACATGCGGCGGAGGTG 481
QY 141 GlnAlaMetLeuGlnLysInsThrGlnGlnLeuArgValArgPheAlaSerHisLeuArg 160
Db 482 CAGGCCATGCTGCGGACAGAGACCGAGAGCTGCGGCTGCGCTCCCTCCACCTGCGG 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180

Db 542 AACCTGGTAAAGCGGCTCCCTCCGATCCGATGACCTGAGAGCGCTGGCATGTAC 601
QY 181 GlnAlaGlnAlaArgGlnGlnGlnAlaGlnArgLysLeuSerAlaThrArgGlnGlnGly 200
Db 602 CAGGCGGCGGCGGCGGAGAGCGGCGGAGCGGCTGACGCGCATCCCGAGCGGCTGGG 661
QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
Db 662 CCCCTGTGTGAACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 721
QY 221 LeuGlnGlnArgAlaGlnAlaThrPgiLysArgLeuArgAlaArgMetGlnGlnMetGly 240
Db 722 CTACAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 781
QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
Db 782 AGTCGAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 841
QY 261 LeuGlnGlnGlnAlaGlnGlnLysArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 842 CTGAGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 901
QY 281 SerTrpPheGlnLysProLeuValGlnAspMetGlnArgGlnThrPheAlaGlyLeuValGlnLys 300
Db 902 AGCTGTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 961
QY 301 ValGlnAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGCAGGCTGCGTGGGACACGCGCCCTGTGCTGCACGACATCTAC 1012

RESULT 11
AX409597 1157 bp DNA linear PAT 14-JUN-2002
LOCUS AX409597
DEFINITION Sequence 2244 from Patent WO0229103.
ACCESSION AX409597
VERSION AX409597.1 GI:21442302
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1
Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2244 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
source location/Qualifiers
1..1157
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M12529"
BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN
Alignment Scores:
Pred. No.: 6.26e-77 Length: 1157
Score: 1564.00 Matches: 314
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 99.05% Mismatches: 3
Query Match: 98.92% Indels: 0
Gaps: 0
DB: 6
US-09-827-854-14 (1-317) x AX409597 (1-1157)
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAGGTTCTGTGGCTGTGGCTGTGCATCTCTGCGAGATGCCAGGCAAGGTG 121
QY 21 GlnGlnAlaValAlaGlyThrGlnProGlnProGlnLeuArgGlnGlnThrGlnThrPgiIns 40
Db 122 GACCAAGCGGTGAGAGACAGAGCGGAGCCGAGCTGCGCAGAGACCGAGTGCAGAGC 181

OY	41	GlyGlnAargTrpPgluLeuAlaIleuGlilYargPherThrPaspyrYLEuArqTryValClnthr	60
Db	182	GGCCAGCGCTGGGAACGTGGACTGGTGGCTTTGGGATTAACCTGCCTGGGTGCAGACA	241
OY	61	LeuSerGIugluInValGIuGIuLleuLeuSerSergInValIThrGIuGIuLeuArgLa	80
Db	242	CTGTCTGAGCAGAGTCBAGAGAGAGCTGCTCAAGCTGCCAAAGTCACCAGAAMATGAGGGCG	301
OY	101	ThrpProVALalaglUGlUgltHrThrArgLAarGLeuSerLySGluLeuGlnAlaIaglnla	120
Db	362	ACCCCCGTAAGGGAGGAGAACCGGGCACGGCTGTCCAAGGAGCTGCACACGGCGTAGGCC	421
OY	121	ArgLeuGIuALaaSPMeGIuaSpVALrArgLYarGLeuvAlaGIIntYArgGYLUval	140
Db	422	CGGCTGGGCGGGCATGTGAGAGACTGTGTGGCGGCTGTGGTGCACTACCGCGCAGGTG	481
OY	141	GlnAlaMetLeuGIyGINserThrcLIuLleuuarVyaLAArgLeuAlaSerHisLEuarY	160
Db	482	CAGGCATAGCTCGGCGACAGACACCGAGAGAGTGGCGGGTGGCGCTCGCTCCACCTGGGC	541
OY	161	LysLEuarGLySArGLeuleuArGNspAlaaspAspleuGlnLYarGLeualaValaylr	180
Db	542	AAGCTGCTTAAGCGGCTCTCTCCGCGATCCCGATCACTGCAGAAACGGCTGGCATTAC	601
OY	181	GlnAlaGIyALArGrGLUGlVALaglUarGrGLyeuSerAlaIEarIGluuArgleuGly	200
Db	602	CAGGCCGGGGCCCCGGAAGGGCGCCAGAGCGGCTCTCACAGGCCATCGGAGCGCTGGGG	661
OY	201	ProLeuVALGIuGIuNGlyArGrVALrArgLAalAthrVALGIySerLEuAlaGIyLnPro	220
Db	662	CCCCTGGTGAACAGGGCGGGTGGGGCGCGCACTGGGGGTCTCCCTGGCGCGCAGCGG	721
OY	221	LeuGIuGIuArGLaIaGIuAlaITrPclYcLIuArGLeuaArGLaIArGMetGIuIMetGLy	240
Db	722	CTACAGAGGCGGGCCAGCGCTGGGGSCAGAGGGCTGC CGCGCATGGAGAGATGGGGC	781
OY	241	SerArqTrHraASpaPARJLeuaspELuValLYsgIuGIuInValaIaGLuValArGLaIlys	260
Db	782	AGTCGACCCCGCACCGGCTGTGACGAGTGAAGGACAGGTGGGAGGTGGCGGCCACAG	841
OY	261	LeuGIuGIuGIuAlaGIuGIuINILearGLeugInlaIaGLaIaphEGlnAlaArgLeuIys	280
Db	842	CTGAGAGGAGCAGGGCCAGCAGATACGCTGCAGCGGAGGGCTTCACAGCCCGCTCAG	901
OY	281	SerTrpPheGIuPrOLEuValGIuaspMeGLnaArgINTrrPaGLyLEuValGIuIys	300
Db	902	AGCTGGTTCGAGCCCTGTGTGAAGACATGTGACAGCGCCAGTGGGGCGGCTGTGGAGAAG	961
OY	301	ValGlnAlaAlaValaGIlyThrSerAlaIaIaProVALrProSerASPasnHis	317
Db	962	GTGCAGGCTGCCTGTGGGCACACGCGCCCGCTGTGTCCAGCAGACAATTCAC	1012
RESULT 12			
LOCUS	115975	1157 bp	DNA linear PAT 03-APR-1996
DEFINITION	Sequence 1 from patent US 5472858.		
ACCESSION	115975		
VERSION	115975.1	GI:1250883	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1157)		
AUTHORS	Attie,A.D., Gretch,D.G., Sturley,S.L. and Beckage,N.E.		
TITLE	Production of recombinant proteins in insect larvae		
JOURNAL	Patent: US 5472858-A 1 05-DEC-1995;		
FEATURES	Location/Qualifiers		
source	I..1157 /organism="unknown"		

BASE COUNT	212 a	370 c	426 g	149 t	ORIGIN
Alignment Scores:					
Pred. No.:	6,26e-77	Length:	1,157		
Score:	1564.00	Matches:	314		
Percent Similarity:	99.05%	Conservative:	0		
Best Local Similarity:	99.05%	Mismatches:	3		
Query Match:	98.92%	Indels:	0		
DB:	6	Gaps:	0		
US-09-827-854-14 (1-317) x IL15975 (1-1157)					
QY	1	MellysValleuTTPpAlaaleuLeuValThPheleuAaelysglnalalySval	20		
Db	62	ATGAAGGCTTCGTGGGGCTGCTCTGTGTACATCTCTGGCAGAGTCCAGGCGAAGG	121		
QY	21	GIuGlnAlaValAlGIuThrGIuPrGIuPrGIuLeuArgGlnGlnThrGIuTPGlnSer	40		
Db	122	GAGCAGACGGGTGGAGACAGAGCCCGGAGCTGGCGCAGCAGACAGAGTGGCAGAC	181		
QY	41	glycInaArTPGIuLeuAlaLeuGIYAArgPheTPAsPTyrLeuArTPYAlGlnThr	60		
Db	182	GGCCAGCGCTCGGAAGTGGACACTGGGTCGCTTTGGATTAACCTGGCGTGGGTGCAGACA	241		
QY	61	LeuSerGIuGlnValGlnGlnGlnLeuLeuLeuSerSerGIuValThrGlnGlnuLeuAga	80		
Db	242	CTGTCTGGAGGAGGAGCAGAGAGACTCTCAAGCTCCCAATCCCAAGAACTGAGGGCG	301		
QY	81	LeuMeTAsPGIuThrMeTlysGIuLeuLysAlaTyrLysSerGIuLeuGlnGlnLeu	100		
Db	302	CTGATGGAGCAGACCATTAAGAGATTGAAGCCCTACAAATCGGAAGCTGGAGGAACTAG	361		
QY	101	ThrProValAlaGIuGlnThrArGIaAArgLeuSerLysGIuLeuGlnAlaAlaGlnAla	120		
Db	362	ACCCCGGTAGCGGAGGAGACCGCGGCAAGCTGTCCAAAGAGCTGCAGACGCGCGAGGCC	421		
QY	121	ArgLeuGlyAlaAsPMeGIuAsPValArgGIYAArgLeuValGIuInYrArgGIuVal	140		
Db	422	CGGGCGGGCGGAGCAATGAAGAGTGTGGCGCGCCCTGTGTGACGTACCGCGCGAGGTG	481		
QY	141	GlnAlaMeTleuGlnGlnSerThrGluGIuLeuArGIaAArgLeuAlaSerHisLeuArg	160		
Db	482	CAGGCATGCTCGGGCAGACAGCAGAGAGAGCTGGGGTGCGCTCGGCTCCACATCGCG	541		
QY	161	LysLeuArGIYSAArgLeuLeuArAsPAlaAsPAsPleuGlnYSAArgLeuAlaValTyr	180		
Db	542	AAGCGCTTAAGCGCGCTCTCCGGAATCCGATGACCTTCAGAAAGCGCCGCGCAAGTGTAC	601		
QY	181	GlnAlaGlyAlaArgGIuGlyAlaGIuArGIYLeuSerAlaLeuArgGIuArgLeuGly	200		
Db	602	CAGGCCGGGGGCCCGGAGGGCGCGGACCGGGCTCAAGCCCTCCGCGAGCGCTGGGG	661		
QY	201	ProLeuValGIuGlnGlnArgValArgAlaAlaThrValGIYSerLeuAlaGlyInPro	220		
Db	662	CCCGTGGGAAACAGGGGGCGGTGCGGGCGCCACCTGTGGGTCCCTGGGGCGGCGCGG	721		
QY	221	LeuGlnGlnArGIaGlnAlaATTPGIYGIuArGIuLeuArGIaAArgMeTGIuMeTGIY	240		
Db	722	CTACAGGAGCGGGCCACAGGCTGGGGCGAGCGCTGGCGCGCGGATGGAGGATGGGCG	781		
QY	241	SerArTPThrAsPArGIaAsPArgLeuAsPGIuValLysGIuGlnValAlaGIuValArGIaLys	260		
Db	782	AGTCGGAGCCCGCGACCGCTTGAGAGAGGTAAAGAGCAGGTGGCGGAGGTGCCCGCAAG	841		
QY	261	LeuGlnGlnGlnAlaGlnGlnInIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys	280		
Db	842	CTGGAGGAGCGGCCCGACAGATCGGCTCCAGAGCCGAGAGGCTTCCAGAGCCGCGCTCAAG	901		
QY	281	SerTPPheGIuProLeuValGIuAsPMeGlnArGIuInTPAlaGIuValGIuLys	300		
Db	902	AGCTGGTTCGAGCCCGGTGGGAAGATACACCGCACAGTGGGGCGCGGTGTGGGGAAG	961		

cy	301	ValGlnAlaAlaValGlyThrSerLalaIapProserAspAsnHis	317
db	962	GTGCAGCTCGCTGGGACACAGCCGCCCTGTGCCACGACATAC	1012
RESULT	13		
LOCUS	HUMAPOE	1157 bp	linear
DEFINITION	Human apolipoprotein E mRNA, complete cds.		PRI 08-AUG-1995
ACCESSION	M12529		
VERSION	M12529.1	GI:178848	
KEYWORDS	apolipoprotein.		
SOURCE	Homio sapiens (clone: PHAE[112,178,813].)	male 57-year old liver	
ORGANISM	Homio sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	McLean,J.W., Elshourbagy,N.A., Chang,D.J., Mahley,R.W. and Taylor,J.M.		
TITLE	Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing of a new variant		
JOURNAL	J. Biol. Chem.	259 (10), 6498-6504	(1984)
MEDLINE	84212473		
PUBMED	6327682		
FEATURES			
Source	Location/Qualifiers		
	1..1157		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/map="19q13.2"		
	/clone="PHAE[112,178,813]."		
	/sex="male"		
	/tissue_type="liver"		
	/dev_stage="57-year old"		
	1..1157		
	/gene="APOE"		
	<1..1157		
	/gene="APOE"		
	/note="apoE mRNA; G00-119-691"		
	62..1015		
	/gene="APOE"		
	/note="precursor"		
	/codon_start=1		
	/product="apolipoprotein E"		
	/protein_id="AAB59518.1"		
	/db_xref="GI:178849"		
	/db_xref="GDB:G00-119-691"		
	/translation="MKVLMALVTPLAGCOAKVEAVETPEPELROPTMGSGGRMR		
	ELAGRPMDYLRWQVTSLEQVQBELLSQVQELRALMDERMKLKAKSKLEQLP		
	VAETRRRLSELDITAOARLDADDEDVCGRLVQTRGEVQALGQSTELRRLASHLR		
	KLRRLRLDPDLDLQRLAVYQAGAREGAERGLSIRRLGLVLEGVQRAVATVSLAG		
	QPLQERAOAMGERLRARMEEGSRTRDLDEVKQVAVEVRAKLEEQAOQIRLQAEAFQ		
	ARLSWFEPLIVEDQROWAGLVEKVAAGVTSAPVPSDNH"		
	62..115		
	/gene="APOE"		
	/note="G00-119-691"		
	116..1012		
	/gene="APOE"		
	/product="apolipoprotein E"		
	/note="G00-119-691"		
	370		
	/gene="APOE"		
	/note="a in PHAE[112,178]; g in PHAE813; G00-119-691"		
	/replace="g"		
	410		
	/gene="APOE"		
	/note="a in PHAE[112,178]; g in PHAE813; G00-119-691"		
	/replace="g"		
	569		
	/gene="APOE"		
	/note="c in PHAE[112,178]; g in PHAE813; G00-119-691"		
	/replace="g"		
	784		
	/gene="APOE"		

variation	/note="t in PHAE[112,178]; c in PHAE813; G00-119-691" /replace="c" 859 /gene="APOE" /note="g in PHAE[112,178]; a in PHAE813; G00-119-691" /replace="a"	
BASE COUNT	212 a	370 c
ORIGIN	30 bp upstream of BstNI site.	426 g
		149 t
Alignment Scores:		
Score:	6.26e-77	1157
Percent Similarity:	1564.00	314
Best Local Similarity:	99.05%	Conservative: 0
Query Match:	98.92%	Mismatches: 3
DB:	9	Indels: 0
		Gaps: 0
US-09-827-854-14 (1-317) x HUMAPOE (1-1157)		
0y	1 MettysValLeutrrPaAlaAlaLeuLeuValThrpheLeuAlaCtysGlnAlaLysVal	20
Db	62 ATGAAGGTTCTGTGGGCTGCTTCTGTGCATTTCTTCGACAGATGCCAGCCAAAGTGTG	121
0y	21 GluGlnAlaValAlaGluThnGluProGluProGluLeuArgGlnGlnThrGluTrrPglInsrt	40
Db	122 GAGCAAGGGGTGGAGACAGACCCGCGACCCAGCTGGCGCAGCAGACCGAGGTGGCAAGC	181
0y	41 GlyGlnArGTrPglLyuAlaAlaLeuGlyArGpThrPsrPTrLeuArGTrPvalGlnThr	60
Db	182 GGCACGCGCTGGGAAGTGGCACTGGTGTGCTTTGGGATTACCTGGCGCTGGTGCACACA	241
0y	61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrnGluLeuArgAla	80
Db	242 CTGTCTGACGAGGTGGACGAGAGAGACTGTCTCAGCTCCCAAGTACCCCAAGAACTGAGGGCG	301
0y	81 LeuMetAspGluTrrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnGlnLeu	100
Db	302 CTGATGGACGAGACCATTAAGACTTGAAGGCTTACAAATTCGGAACCTGGAGAAACAATG	361
0y	101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	362 ACCCGGTAAGCGAGAGACCGCGGGCAAGCTGTCCAAAGAGCTGCAGACGCGCGACGCC	421
0y	121 ArgLeuGlnAlaAspMetGluAspValArgGlyArgLeuValGlnTrArgGlyGluVal	140
Db	422 CGGCTGGGCGCGACATGAGAGAGAGCTGGCGCGCTGTGGCAGTACCAGCGGAGGTG	481
0y	141 GlnAlaMetLeuGlyGlnSerThnGlnGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	482 CAGGCAATGCTCGGCCAGAGACACCGAGAGACTGGGGTGGCGCTCGCTCCACATGCGC	541
0y	161 LysLeuArGTrAsrGLeuLeuArAspAlaAspAspLeuGlnLysArGLeuAlaValAlaTr	180
Db	542 AAGCTGCGTAAAGCGGCTCTCCGCGATCCCGATGACCTGCAGAAAGCGCTGGCAGTGTAC	601
0y	181 GlnAlaGlyAlaArGLeuGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200
Db	602 CAGGCGGGGCGCGCGAGAGCGCGGCTCTCAGCCGCAATCCGCGAGCGCGCTGGCG	661
0y	201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro	220
Db	662 CCCCTGGTGAACAAGGCGGGGTGGGCGCGCCACTGTGGGCTCCCTGGCGGCGCACCG	721
0y	221 LeuGlnGluArGTrAglnAlaTrPrgLysGluArgLeuArgAlaArgMetGluGlnMetGly	240
Db	722 CTACAGAGCGCGCGCCAGGCTGGGGCGAGAGCGCTGGCGCGCGCGAGTGAAGATGGCC	781
0y	241 SerArGTrHArGAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys	260
Db	782 AGTGGAGACCGCGACCGGCTGAGAGGTGAAGAGCAGGATGGGGAGTGGCGGCCAAG	841
0y	261 LeuGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys	280

Db 842 CTGGAGAGCAGCCAGCCAGCATATGCGCTTCAGAGCCCTTCAGGCCCTTCAG 901
Oy 281 SerTrrPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluIys 300
Db 902 AGCTGTTGAGACCCCTGTTGAGACATGACGCGCATGGCGGCTGTTGAGAG 961
Oy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGCAGAGCTGCGTGGGACACAGCGCGCCCTGTGGCCAGGACATTCAC 1012
RESULT 14
MFAPOE 1178 bp mRNA linear PRI 31-MAR-1995
LOCUS Monkey mRNA for apolipoprotein E.
DEFINITION X13887
ACCESSION X13887
VERSION X13887.1 GI:38054
KEYWORDS apolipoprotein; apolipoprotein E.
SOURCE Macaca fascicularis.
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 1178)
AUTHORS Marotti K.R.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1988) Marotti K.R., The Upjohn Co, 7242 209 713,
301 Henrietta Street, Kalamazoo, MI 49008
2 (bases 1 to 1178)
Marotti K.R., Whitted B.E., Castle C.K., Polites H.G. and
Meichlor G.W.
TITLE Nucleotide sequence of the cynomolgus monkey apolipoprotein E cDNA
JOURNAL Nucleic Acids Res. 17 (4), 1778 (1989)
MEDLINE 89160349
PubMed 2923300
COMMENT Data kindly reviewed (20-Mar-1989) by Marotti K.R.
FEATURES
source
location/Qualifiers
1..1178
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="1"
/cissue_type="liver"
/clone_lib="lambda gt11"
83..1036
/codon_start=1
/product="proapolipoprotein E (AA -18 to 299)"
/protein_id="CAA32092.1"
/db_xref="GI:38054"
/db_xref="SWISS-PROT:P10517"
/translation="MKVLMALVTFPLAGCOAKVEOPVEPEPELROOAGSGGQPM
ELALGKEMDYLRMVOTLSEVOEELISPOVQOELTTIMDETMRKELKAYKSLLEOLSP
VAETRRARLSKELOAAOARLADMEDEVSRVYORSEVOAMLGSTELRRLASHLR
KLKRLDLADDDQLRLAVYQAGABEGEVSALIRELGLVBOGVRATVGSLSL
QPLDERAOLAEGLRLARMEGMSRTKRLDEKQVAVRAKLEEOAQOISLQAEARQ
ARLKSWEPLVEDMQRQWAGLVEKVOAAVASTAPVIDNH"
83..136
sig_peptide
mat_peptide
misc_feature
137..1033
/product="apolipoprotein E (AA 1 - 299)"
1154..1159
/note="polya signal"
BASE COUNT 217 a 373 c 426 g 162 t
ORIGIN

Alignment Scores:
Pred. No.: 3,17e-71 Length: 1178
Score: 1459.00 Matches: 297
Percent Similarity: 94.32% Conservative: 2
Best Local Similarity: 93.69% Mismatches: 18
Query Match: 92.28% Indels: 0
DB: 9 Gaps: 0

US-09-827-854-14 (1-317) x MFAPOE (1-1178)
Oy 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20

Db 83 ATGAAGGTTCTGTGGGCTGCGTGTGTCATATTCGTGGAGATGACAGGCAAGGTG 142
Oy 21 GluGlnAlaValGluThrGluProGluProGluLeuAlaGlnGlnThrGluTrpGlnSer 40
Db 143 GAGCAACCGGTGGAGCCAGACAGACGCCGAGCTTGGCCAGACGATGAGGCCAGAC 202
Oy 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpAlaGlnThr 60
Db 203 GGCCAGCCCTGGGAGCTGGACCTGGACTGGGTGCTTTTGGATTAAGCTGCTGGGTGACACA 262
Oy 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db 263 CTGTGTAGACAGGTGACAGAGAGAGCTGCTCAGCCCCCAGAGTCACCCAGGAAGTACAGAC 322
Oy 81 LeuMetAspGluThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
Db 323 CTGATGGACGAGACCATGAAAGAGCTTGAAGGCTCAATCGGAATCGGAGACACAGCTG 382
Oy 101 ThrProValAlaGlnGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
Db 383 AGCCCGGTGGGAGAGACGCGGCGACGCTGTCAAAGAGCTGACAGCGGCGGCAAGGCC 442
Oy 121 ArgLeuGlnAlaAspMetGlnAspValArgGlyArgLeuValGlnTrpArgGlyGlyVal 140
Db 443 CGGCTGGGTGGCGACATGAGAGAGAGCTGGCGGACCGCTGTGTCACTACCGCAGGAGGTG 502
Oy 141 GlnAlaMetLeuGlyGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
Db 503 CAGGCGATGCTGGGCGCAAGATGACAGAGAGCTGGCGGCGGCGCTGCCCTCCACCTGGCG 562
Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 563 AAGCTGCGGACAGCGGCTCTCCGCGATGCTGATGACTGACAGAGAGCGCTTGCACTGTAT 622
Oy 181 GlnAlaGlyAlaArgGlyGlnGlyValGlnGlnArgGlyLeuSerAlaIleArgGlnArgGly 200
Db 623 CAGCGCGGCGGCGCGGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 682
Oy 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 683 CCCCTGGTGGAGCAGGCGCGGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 742
Oy 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
Db 743 CTTCAGGAGCGGCGCGGAGGCGCTTGGGTAGCGGCTTCCGCGACGGATGGAGAGATGGGC 802
Oy 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValAlaArgAlaLys 260
Db 803 AGCCGAGACCCCGGACCGGCTGGAGCAGGTGAAGAGACAGGTGGCGAGGTGCGGCCCAAG 862
Oy 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 863 CTGCAGGAACAGCGCCACAGATAGCTTCGAGGCGGAGGCGCTTCCAGCGCGGCTCAAG 922
Oy 281 SerTrrPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGluIys 300
Db 923 AGCTGTTGAGACCCCTGTTGAGACATGACGCGCATGGCGGCTGTTGAGAGAG 982
Oy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 983 GTGCAGAGCTGCGTGGGACACAGCGCGCCCTGTGGCCAGGACATTCAC 1012
RESULT 15
HUMAPOE4 5515 bp DNA linear PRI 09-NOV-1994
LOCUS Human apolipoprotein E (epsilon-4 allele) gene, complete cds.
DEFINITION M10065 J03053 J03054
ACCESSION M10065.1 GI:178852
VERSION M10065.1 GI:178852
KEYWORDS Alt repeat; allelic variation; apolipoprotein; apolipoprotein E;
lipoprotein; repeat region; very low density lipoprotein.
SOURCE Human DNA [2], [1].
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 5515)
Das,H.K., McPherson,J., Bruns,G.A., Karathanasis,S.K. and Breslow,J.L.
Isolation, characterization, and mapping to chromosome 19 of the human apolipoprotein E gene
J. Biol. Chem. 260 (10), 6240-6247 (1985)
MEDLINE 85207610
PubMed 3922972
2 (bases 196 to 5269)
Paik,Y.K., Chang,D.-J., Reardon,C.A., Davies,G.E., Mahley,R.W. and Taylor,J.M.
Nucleotide sequence and structure of the human apolipoprotein E gene
Proc. Natl. Acad. Sci. U.S.A. 82 (10), 3445-3449 (1985)
JOURNAL 85216517
MEDLINE 2987927
PubMed 3 (bases 1 to 5515)
Williams,R.R., White,R. and Lalouel,J.M.
Genotyping and sequence analysis of apolipoprotein E isoforms
Genomics 3 (4), 373-379 (1988)
JOURNAL 89212602
MEDLINE 3243553
PubMed
[3] two allelic variations.
Draft entry and computer-readable sequence for [3] kindly provided by M.Emi, 19-AUG-1988.
Apolipoprotein E is a constituent of the human very low density lipoprotein in the plasma. There are at least six distinct phenotypes derived from the single E gene on chromosome 19; next to the epsilon-3 allele (see separate entry), the epsilon-4 allele, represented by the sequence below, is most common, the product difference being arginine in place of cysteine at residue 112 [2]. The gene structure of apo E is similar to that of other apo genes: presence of the 66-bp repeats in the fourth exon (starting at base 3782 below) makes the E gene highly similar to the A-I gene (see separate entry) as argued by [1].
A potential TATA box is found at positions 1014-1018, and a potential polyadenylation signal at 4616-4621.
[2] and [1] had slight differences in the boundary positions for the Alu repeats and their flanks: the boundary positions indicated in [1] have been used in the FEATURES table below. Draft entries and clean copies were kindly supplied by J.M. Taylor, Gladstone Laboratories, San Francisco, and by J.P. Levine, Rockefeller University, New York.
Location/Qualifiers
1..5515
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19q13.2"
202..203
/citation=[2]
/replace=""
repeat_region
/note="direct repeat flanking Alu repeat 3' copy [1]"
complement(346..635)
/note="Alu repeat [1]"
repeat_region
complement(636..680)
/note="direct repeat flanking Alu repeat 5' copy [1]"
963..967
/citation=[2]
/replace=""
exon
1047..1090
/note="apo E mRNA [2],[1]"
/number=1
1091..1847
/note="apo E mRNA intron 1 [2],[1]"
1362..1363
/citation=[2]
/replace=""
1789..1793
/citation=[2]

CDS
/replace=""
join(1871..1913,3007..3199,3781..4498)
/note="precursor"
/codon_start=1
/product="apolipoprotein E"
/protein_id="AA059397.1"
/db_xref="GI:178853"
/translation="MKVLAALLVFLLACQAKVEQAVETEPPELROQTEWQSGQRM ELAIGFMDYLRWVOTLSQVOBELSSQVTOELRALMDETKELKAVKSELEOLTP VAETFRARLSKETLAAQARLQADIMEDYRGRGLVQYRGEOAMLGOSTEELRVLASHLR KLRRLRQADIDQRLAYVQAGAREGRLSAIERLGRPLVEQGRVAARVAVGSIAG QPLQERQAQWGCRLRRLRMEGSRTPEDRLDEYAEVRAKLEDAQQIRLQAEAFQ ARKSWFEPLVEDMQRMAGLVKQVAAVGTSAAPVSDNH"
join(1871..1913,3007..3017)
join(3018..3199,3781..4495)
/product="apolipoprotein E"
1871..1913
/gene="APOE"
1871..1913
/partial
/gene="APOE"
/note="preapolipoprotein E; G00-119-691"
/number=2
1914..3006
/note="apo E cds intron 2 [2],[1]"
complement(2092..2104)
/note="direct repeat flanking Alu repeat 3' copy [1]"
complement(2105..2429)
/note="Alu repeat [1]"
complement(2430..2442)
/note="direct repeat flanking Alu repeat 5' copy [1]"
2520..2526
/note="direct repeat flanking Alu repeat 5' copy [1]"
2527..2886
/note="Alu repeat [1]"
2568..2569
/citation=[2]
/replace=""
repeat_region
2887..2893
/note="direct repeat flanking Alu repeat 3' copy [1]"
2947..2948
/citation=[2]
/replace=""
conflict
2983..2984
/citation=[2]
/replace=""
exon
3007..3199
/number=2
3182
/note="g in epsilon-4; a in epsilon-3 [2]"
3191
/note="g in epsilon-4; a in epsilon-3 [2]"
3200..3780
/note="apo E cds intron 3 [2],[1]"
3568..3569
/citation=[2]
/replace=""
exon
3781..4640
/partial
/note="preapolipoprotein E"
/number=3
3817..3819
/note="cta in [2] and 1 allele [3]; agc in other allele [3]"
3844..3846
/note="gac in [2] and 1 allele [3]; tgt in other allele [3]"
3932
/note="c in epsilon-4; t in epsilon-3 [2] (arg; cys)"
4342
/note="g in epsilon-4; a in epsilon-3 [2]"
4708..4709
/citation=[2]


```

/replacement="
repeat_region complement(4761..4768)
/notes="direct repeat flanking Alu repeat 3' copy [1]"
repeat_region complement(4769..5048)
/notes="Alu repeat [1]"
repeat_region complement(5049..5056)
/notes="direct repeat flanking Alu repeat 5' copy [1]"
BASE COUNT 1042 a 1667 c 1600 g 1206 t
ORIGIN 201 bp upstream of BanII site on chromosome 19q12-q13.2.

Alignment Scores:
Pred. No.: 1.24e-66 Length: 5515
Score: 1387.00 Matches: 307
Percent Similarity: 58.22% Conservative: 1
Best Local Similarity: 58.03% Mismatches: 8
Query Match: 87.73% Indels: 214
DB: Gaps: 2

US-09-827-854-14 (1-317) x HUMAPOA4 (1-5515)
Qy 2 LysValLeuTPpAlaAlaLeuVal----- 10
Db 2910 AAGCATTGTGTGGAGCAGCTTCTGTGTGCTAGCTAGCTAGATGCTGGACGGGGCTAG 2969
Qy 11 -----ThrpheLeuAlaGlyCysGlnAlaLysValGln 22
Db 2970 AAGGACCTGACCCGACCTTGAAGCTGTTCACACAGATGCCAGGCCAGAGTGAGCA 3029
Qy 23 AlaValGlnThrGluProGluProGluLeuArgGlnThrGluThrGlnSerGlyGln 42
Db 3020 GCGGGGAGACAGAGCCGAGCCGAGCTGCGCGCCAGACGAGAGTGAGCAGAGGGCCAG 3089
Qy 43 ArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThrLeuSer 62
Db 3090 GCGTGGAGACTGGACAGCTGGGCTTTGGGATTACCTGGCTGGCTGACAGACTGTCT 3149
Qy 63 GluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeu----- 78
Db 3150 GAGCAGGTGACAGAGAGACTGCTCAGCTCCAGGTACCCAGAACT-GAGGTAGTGTCTC 3208
Qy 78 ----- 78
Db 3209 CCCATCTCGCCCTTGACCTCTCTGTGGGGCTATACCTCCAGAGTCCAGGTTTCAT 3268
Qy 78 ----- 78
Db 3269 TCTGCCCCCTTGCGCTAAGCTTTGGGGGGCTGGGCTCTCTGCTGTAGCTTCTCTTC 3328
Qy 78 ----- 78
Db 3329 CCATTTCGACTCGCGCTTAAAGCTCTGGAATTCTCTCTGACGTTTGTCTCTCTCT 3388
Qy 78 ----- 78
Db 3389 CTTCCTTCTGACTCAGTCTCTCACACTCTGCTGTCTGTCTGTCTCTCTCTCTCT 3448
Qy 78 ----- 78
Db 3449 CTTTATATAGAGACAGAGAGATGGGGTCTCAGTGTGTGGCCAGGCTGGTCTGAAGCTT 3508
Qy 78 ----- 78
Db 3509 CTGGGCTCAAGCAGATCTCTCCGGCTCGCCCAAGTGTGGGATTAGAGGATGAGC 3568
Qy 78 ----- 78
Db 3569 ACCTTGGCCGGGCTCTAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3628
Qy 78 ----- 78
Db 3629 TGCATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3688
Qy 78 ----- 78

```

```

Db 3689 CTCCTGCGTCATCCCATCTCGCCGCCCATCCAGCCCTTCTCCCCCGCTCCAC 3748
Qy 79 -----ArgAlaLeuMetAspGlnThrMetLysGlu 88
Db 3749 TGTGGACACCCCTCCGCCCTCTCGCGCCGACGGGGCTGATGGAGACCATGTAAGAG 3808
Qy 89 LeuLysAlaTrpLysSerGlnLeuGlnGlnLeuThrProValAlaGlnGlnThrArg 108
Db 3809 TTGAAGGCTTACAAATCGGAAGTGGAGAACACTGACCCCGGTGGCGAGAGCGCGG 3868
Qy 109 AlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAlaArgLeuGlnAlaAspMetGluAsp 128
Db 3869 GCACGCGTGTCCAAAGACACTGACAGCGCGGACGGCCCGGTGGCGCGACATGAGAGAC 3928
Qy 129 ValArgGlyArgLeuValGlnThrArgGlyGlnValGlnAlaMetLeuGlyGlnSerThr 148
Db 3929 GTGGCGGGCCGCTGTGCTACAGTACCGCGGAGGTGACAGGCCATGCTGCGCCAGACAC 3988
Qy 149 GluGlnLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuValArg 168
Db 3989 GAGGAGCTGCGGGTGCCTGCTGCTCCCACTGCGCAAGCTGCTAAGCGGCTCTCTCCGC 4048
Qy 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTrpGlnAlaGlyAlaArgGlnGlnVal 188
Db 4049 GATGCCGATGACCTGACAGAGCGCTGCGCAGTACAGGCCGGGGCCGAGAGGGCGCC 4108
Qy 189 GluArgGlyLeuSerAlaIleArgGlnArgLeuGlyProLeuValGlnGlnArgVal 208
Db 4109 GAGCGCGGCTTCAGCGCCATCGCGGAGCGCTGGGGGCCCTGTGTGGAAACAGAGCGCGCTG 4168
Qy 209 ArgAlaIleThrValGlySerLeuAlaGlyGlnProLeuGlnArgAlaGlnAlaTrp 228
Db 4169 GGGGCGGCACTGTGGGCTTCCCTGCGGCCAGCGCTACAGAGAGCGGGCCAGGCTGTG 4228
Qy 229 GlyGlnArgLeuArgAlaArgMetGlnGlnMetGlySerArgTrpArgAspArgLeuAsp 248
Db 4229 GCGCAGCGCGCTGCGCGCGCGGATGAGAGATGGGACCGGACCGGACCGCTGTGAC 4288
Qy 249 GluValLysGlnGlnValAlaGlnValArgAlaLysLeuGlnGlnAlaGlnIle 268
Db 4289 GAGGTGAAGAGACAGGTGGCGGAGGTGCGCCAAAGCTGGAGAGACAGCCAGCAGATA 4348
Qy 269 ArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLysSerTrpPheGlnProLeuValGlu 288
Db 4349 GCGCTGACAGCGCGAGGCTTCCAGGCGCGCTCAAGAGCTGTGTGAGCCCTGTGGAA 4408
Qy 289 AspMetGlnArgGlnTrpAlaGlyLeuValGlnLysValGlnAlaAlaValGlyThrSer 308
Db 4409 GACATGACAGCGCCAGGTGGCGCGGCTGTGGAGAAAGTGCAGAGGTGCGCTGGGACACAGC 4468
Qy 309 AlaAlaProValProSerAspAsnHis 317
Db 4469 GCCGCCCTGTGCCAGCGCAATTCAC 4495

```

Search completed: March 14, 2003, 17:32:05
 Job time : 2316.23 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:08:17 ; Search time 178.728 seconds

(without alignments)
3994.237 Million cell updates/sec

Title: US-09-827-854-14

Perfect score: 1581

Sequence: 1 MKVLMAALVFLAGCOAKV.....VEKYQAVGTSAPVPSDNH 317

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlp
-O=/cgn2.1/USPFO.spool/US09827874/runit_11032003.101609.27476/app_query.fasta.1.3576
-DB=N.Geneseq.101002 -QFMT=fastlap -SUFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum2 -TRANS=human40.cdl
-LIST=45 -DOCCALIGN=200 -THR.score=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09827874.ecgn.1.1.1201 -runatc_11032003.101609.27476 -NCPU=6 -ICPU=3
-NO_XLIPX -NO_WMAP -LARGEQUERY -NEGSCORES=0 -WAIT -LONGLOGS -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELExt=7

Database :
1: N.Geneseq.101002.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
25: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1581	100.0	1156	22 AAF84316	Human ApoE4 coding
2	1581	100.0	1156	24 AAD22047	Human apolipoprote
3	1573	99.5	954	24 AAD26035	Human apolipoprote
4	1573	99.5	1110	7 AAN60409	Human apolipoprote
5	1573	99.5	1147	22 ABA83113	Apolipoprotein E o
6	1573	99.5	1156	22 AAF84315	Human ApoE3 coding
7	1573	99.5	1156	24 AAD22048	Human apolipoprote
8	1569	99.2	1156	24 AAD22052	Human apolipoprote
9	1565	99.0	1156	22 AAF84314	Human ApoE2 coding
10	1565	99.0	1156	24 AAD22049	Human apolipoprote
11	1565	99.0	1156	24 AAD22051	Human apolipoprote
12	1564	98.9	1157	17 AAT06957	Human apolipoprote
13	1564	98.9	1157	24 ABN95746	Gene #2244 used to
14	1564	98.9	1157	24 ABK64514	Human benign prost
15	1564	98.9	1157	24 ABL65450	Lung cancer relate
16	1563	98.9	1110	6 AAN50450	Sequence encoding
17	1558	98.5	1156	24 AAD22050	Human apolipoprote
18	1469.5	92.9	1279	22 AAS22437	Human CDNA encodin
19	1450.5	91.7	1107	19 AAX75756	Human apolipoprote
20	1384.5	87.6	9360	24 ABL31915	Human ApoE genomic
21	1384.5	87.6	10716	24 AAD26034	Human apolipoprote
22	1379	87.2	3805	20 AAZ09524	Human Apo E genom
23	1379	87.2	3805	20 AAZ09526	Human Apo E genom
24	1370.5	86.7	10716	24 AAD26108	Human apolipoprote
25	1151.5	72.8	965	24 AAD32081	Bovine ApoE gene.
26	1151.5	72.8	5617	24 AAD32077	Human alpha-1-anti
27	1151.5	72.8	6026	24 AAD32075	Human albumin prom
28	1146	72.5	1126	19 AAV29159	Nucleotide sequenc
29	993	62.8	936	15 AAO69101	ApoE4Lx2 protease
30	993	62.8	936	17 AAT18070	Human ApoE4Lx2 CDN
31	972	61.5	660	18 AAT69792	Partial human apol
32	965	60.4	597	17 AAT18068	Human ApoE4 CDNA.
33	946	59.8	597	15 AAO69099	ApoE4L protease cd
34	872	55.2	1381	22 AAS22673	Human CDNA encodin
35	867.5	54.9	786	21 AAF18114	Lung cancer associ
36	686	43.4	600	20 AAV89595	EST clone CP147.
37	651	41.2	407	24 AABK3238	Human CDNA for nov
38	625	39.5	478	24 AABN5994	Gene #2492 used to
39	625	39.5	478	24 ABL62679	Colon adenocarcino
40	625	39.5	478	24 ABL67340	Thyroid cancer rel
41	594	37.6	499	22 AAB03049	Human diagnostic a
42	522	33.0	330	12 AAQ11980	Human apolipoprote
43	504	31.9	405	21 AAC02139	Human secreted pro
44	485	30.7	345	22 AAH98479	Human EST-derived
45	410	25.9	382	21 AAA44191	Mouse secreted exp

ALIGNMENTS

RESULT 1
AAE84316
AAE84316 standard; CDNA: 1156 BP.

AC AAE84316;
21-JUN-2001 (first entry)

DE Human ApoE4 coding sequence.

KW Human; ApoE4; Alzheimer's disease; arteriosclerosis; ss.

OS Homo sapiens.

FT key
FT CDS
FT Location/Qualifiers
61..1014
/*tag= a
/product= "Human ApoE4"

XX JP2001017028-A.
 PN
 XX
 PD 23-JAN-2001.
 XX
 PF 28-APR-2000; 2000JP-0128919.
 XX
 PR 06-MAY-1999; 99JP-0125647.
 XX
 PA (MITU) MITSUBISHI CHEM CORP.
 XX
 DR WPI; 2001-285406/30.
 XX P-PSDB; AAB80998.
 XX
 PT New apoE humanized mammalian cell useful for screening for agents
 CC useful for treating or preventing Alzheimer's disease and
 CC arteriosclerosis -
 XX
 PS Disclosure; Page 16-17; 22pp; Japanese.
 XX
 CC The present invention relates to an ApoE humanised mammalian cell. The
 CC present sequence is the coding sequence for human ApoE4, which was used
 CC in the method of the present invention. The ApoE humanised mammalian cell
 CC can be used for screening for agents useful for treating or preventing
 CC Alzheimer's disease and arteriosclerosis.
 XX
 SQ Sequence 1156 BP; 208 A; 369 C; 432 G; 147 T; 0 other;

Alignment Scores:

Pred. No.: 4 64e-112 Length: 1156
 Score: 1581.00 Matches: 317
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-827-854-14 (1-317) x AAF84316 (1-1156)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 61 ATGAAGTTCTGTGGCTGTGCTGTCTGTACATCTCTGGCAGATGCCAGGCCAAGTGT 120
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 121 GAGCAACGGCTGGAGACAGACCGGACCGGAGCTGGCCAGCAGACCGAGTGGCAGAGC 180
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 DB 181 GGGCAGCGCTGGGAACCTGGCACTGGCTGCTTTGGGATTACCTGGCTGGTGCAGACA 240
 QY 61 LeuSerGlnGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB 241 CTGTCTAGCAGAGTGGCAGGAGAGCTGTCTAGCTCCAGGTCACCCAGGAACCTAGGGCG 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGluGlnLeu 100
 DB 301 CTGATGACAGACCAATGAGAGATTGAAGCTTACAAATCGGAACGTGGAACAACCTG 360
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 361 ACCCGGGTGGCGAGAGACGCGGCGGCTGTCCAAAGAGCTGCGAGGCGCGCAGGCC 420
 QY 121 ArgLeuGlnAlaAspMetGluAspValArgGlyArgLeuValGlnTrpArgGlyGluVal 140
 DB 421 CGGCTGGGCGCGGACATGAGAGACGTGGCGGCGCTGTGTCAGTACCGCGGCGAGGTG 480
 QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 481 CAGGCAATGCTCGGACAGACACCGAGAGACTGCGGGTGGCTCGCTCCACACTCGCG 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspSerLeuGlnLysArgLeuAlaValTrp 180
 DB 541 AAGCTCGTAAGCGGCTCTCCGCGATGCCGATGACCTGCAGAAAGCGCCGTGGAGTGTAC 600

QY 181 GlnAlaGlyAlaArgGluGluGluAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 601 CAGCCCGGCGCCCGGAGGCGCGGAGCGGCGCTCAGCCCATCCGAGAGCGCTGGG 660
 QY 201 ProLeuValGluGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 DB 661 CCCCTGGTGGAAACAGGGCCCGCTGGGGCCGCACTGTGGGCTTCCCTGGCGCGGACCCG 720
 QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
 DB 721 CTACAGAGAGCGGCGCCAGGCTTGGGGCGAGCGGCTGGCGCGATGGAGATGGGC 780
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 DB 781 AGCCGAGACCCGCGACCCCTGTGACAGAGTGAAGAGAGAGGTGGCGGAGTGGCGCCAG 840
 QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 DB 841 CTGGAGGAGCAGGCGCCAGCATACGCTTCAGGCGCAGGCTTCCAGGCGCCCTTCAG 900
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 DB 901 AGCTGTTCCAGCCCTGTGGTGAAGACATGACGCGCCAGTGGCGGCTGTGTGAGAG 960
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 961 GTGCAGGCTCGGTGGCACCAGCGCGCCCTGTGCCCCAGCAGCATTCAC 1011

RESULT 2

AAD22047 standard; DNA; 1156 BP.

AAD22047;

12-FEB-2002 (first entry)

Human apolipoprotein E (apoE) isoprotein, apoE4 DNA.

Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;

KM hypertirglyceridaemia; low density lipoprotein; LDL; ds.

Homo sapiens.

Key Location/Qualifiers

FT sig_peptide

FT mat_peptide

FT /tag= a /product= "Human apoE isoprotein, apoE4"

FT /tag= b

FT /tag= c

FT /product= "Mature human apoE isoprotein, apoE4"

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

WO200177136-A1.
 18-OCT-2001.
 PD
 XX
 PD
 XX
 PF 06-APR-2001; 2001MO-US11358.
 XX
 PR 06-APR-2000; 2000US-0544386.
 PR 04-OCT-2000; 2000US-0679088.
 PR 05-APR-2001; 2001US-0827854.
 PA (KOSP-) KOS PHARM INC.
 PA (UYBO-) UNIV BOSTON.
 XX
 PI Zannis VI, Kypros KE;
 XX
 DR WPI; 2002-010885/01.
 XX P-PSDB; AAE13293.
 XX
 PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering

XX

XX

CC the invention.

XX Sequence 1147 BP; 210 A; 365 C; 425 G; 147 T; 0 other;

Alignment Scores:

Pred. No.:	1,88e-111	Length:	1147
Score:	1573.00	Matches:	316
Percent Similarity:	99.68%	Conservative:	0
Best Local Similarity:	99.68%	Mismatches:	1
Query Match:	99.49%	Indels:	0
DB:	22	Gaps:	0

US-09-827-854-14 (1-317) x ABA83113 (1-1147)

```

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 46 ATGAAGTTCTGTGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 105
QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluThr 40
DB 106 GACCAAGCGGTGGAGACAGACCGGACCGACCTGCGCCAGACAGACGAGTGGCAGAGC 165
QY 41 GLYlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 166 GGCACAGCGCTGGAACTGGCACTGGCTGCTTTGGATTACTGCTGGTGGTGGACACA 225
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 226 CTGTCTGACGAGGTGGAGGAGGAGCTGTCAAGCTCCAGGTCAACCCAGGAAGTGGAGCG 285
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
DB 286 CTGATGGACAGACCAAGAGAGTTGAAGCTTACAAATCGGAACCTGGAGGAACAACCTG 345
QY 101 TheProValAlaGluGluThrArgAlaArgLeuSerLysGlnGlnAlaAlaGlnAla 120
DB 346 ACCCGCGTGGCGAGAGACCGCGGCGGCTGTCCAGAGAGTGGCGCGCGAGGCG 405
QY 121 ArgLeuGlnAlaAspMetGlnAspValArgGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 406 CGGCTGGCGCGGACATGAGAGAGCTGTGGCGGCTGTGGAGTACCGCGGAGAGTGT 465
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 466 CAGCGCATGTCTGGCCAGACACCGAGAGACTCGGGTGGCTCGCTCCACCTGCGC 525
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 526 AAGCTGCGTAAGCGGCTCTCCCGCATGCGGATGACCTGAGAAAGCCCTGGCACTTAC 585
QY 181 GlnAlaGlyAlaArgGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 586 CAGCGCGGCGCGCGCGAGCGCGCGAGCGCGGCTCAAGCCATCCGCGAGCGCCCTGGGG 645
QY 201 ProLeuValGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 646 CCCCTGTGGAAAGAGCGCGCTGCGGGCGCGCACTGTGGCTCCCTGGCGCGCGCAGCGC 705
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
DB 706 CTACAGAGAGGCGGCCAGGCTGGGGCGAGCGGCTGCGCGCGGGATGGAGAGATGGGC 765
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAlaArgAlaLys 260
DB 766 AGCGGACCGCGGACCGCGCTGGAGAGGTGAAGAGAGAGTGGCGGAGGTGGCGCGCAAG 825
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
DB 826 CTGGAGAGACAGCGCCAGAGATACCTGACAGCGGAGGCTTCCAGAGCGCCGCTCAAG 885
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 886 AGCTGTGTGAGCCCTGTGTGGAAGACATGACAGCGCTAGTGGCGGGCTGTGTGGAGAGAG 945

```

```

QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 946 GTGCAGGCTGCGCGTGGGACAGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 996

```

RESULT 6

AAAF84315

ID AAAF84315 standard; cDNA; 1156 BP.

AC AAAF84315;

XX 21-JUN-2001 (first entry)

DE Human APOE3 coding sequence.

XX Human; APOE3; Alzheimer's disease; arteriosclerosis; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 61..1014

FT /tag= a

FT /product= "Human APOE3"

PN JP2001017028-A.

XX 23-JAN-2001.

PD 28-APR-2000; 2000JP-0128919.

PF 06-MAY-1999; 99JP-0125647.

XX (MITU) MITSUBISHI CHEM CORP.

XX WPI: 2001-285406/30.

DR P-PSDB; AAB80997.

PT New apoE humanized mammalian cell useful for screening for agents

PT useful for treating or preventing Alzheimer's disease and

PT arteriosclerosis -

XX Disclosure; Page 13-14; 22pp; Japanese.

XX The present invention relates to an APOE humanised mammalian cell. The

XX present sequence is the coding sequence for human APOE3, which was used

XX in the method of the present invention. The APOE humanised mammalian cell

XX CC Alzheimer's disease and arteriosclerosis.

XX Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other;

Alignment Scores:

Pred. No.:	1.9e-111	Length:	1156
Score:	1573.00	Matches:	316
Percent Similarity:	99.68%	Conservative:	0
Best Local Similarity:	99.68%	Mismatches:	1
Query Match:	99.49%	Indels:	0
DB:	22	Gaps:	0

US-09-827-854-14 (1-317) x AAAF84315 (1-1156)

```

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 61 ATGAAGTTCTGTGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluThr 40
DB 121 GACCAAGCGGTGGAGACAGACCGGAGCGGACCTGCGCCAGACAGACGAGTGGCAGAGC 180
QY 41 GLYlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 181 GGCACAGCGCTGGAACTGGCACTGGCTGCTTTGGATTACTGCTGGTGGTGGACACA 240

```


QY 61 LeuSerGIuGIuValGIuGIuLeuSerGIuValThrGIuLeuArgAla 80
 DB 241 CTGTCTGACGAGGTGACAGAGAGCTGCTCCAGCTCCAGAGCACTGAGGCGC 300
 QY 81 LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuLeuLeu 100
 DB 301 CTGATGGACGACCATGAAAGAGTTGAAGGCTTCAAAATCGGAAGCTGAGAGCAACTG 360
 QY 101 ThrProValAlaGIuGIuThrArgAlaArgLeuSerLysGIuLeuGIuAlaGIuAla 120
 DB 361 ACCCGGTGGCGGAGAGACGCGGGCAGCGCTGTCCAAAGAGCTCAGCGCGGCAAGCC 420
 QY 121 ArgLeuGIuAlaAspMetGIuAspValArgGIuArgLeuValGIuTyrArgGIuVal 140
 DB 421 CGGCTGGCGGCGACATGAGAGACGTGTGGCCCGCTGTGTGCTACCGCGGCGAGGTG 480
 QY 141 GIuAlaMetLeuGIuGIuSerThrGIuLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 481 CAGGCCATGCTCGCCACAGACACCGAGAGAGCTGCGGTGCGCTCCACCTGCGC 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGIuLysArgLeuAlaValTyr 180
 DB 541 AAGCTGCTGACGCGCTCTCCGCGATGCCATGACCTGCAAGAGCGCTGGAGTGTAC 600
 QY 181 GIuAlaGIuAlaArgGIuGIuValArgGIuLysSerAlaAlaArgGIuArgLeuGIu 200
 DB 601 CAGCGCGGGCGCGGAGGCGCGCGCGCTCAGCGCATCGCGGAGCGCTGGGG 660
 QY 201 ProLeuValGIuGIuGIuArgValArgAlaAlaThrValGIuSerLeuAlaGIuPro 220
 DB 661 CCCCTGTGTGAAACAGGCGCGCTGCGCGCGCGCTGTGTGCTCCCTGCGGCGAGCGG 720
 QY 221 LeuGIuGIuArgAlaGIuAlaThrGIuArgLeuArgAlaArgMetGIuGIuMetGIu 240
 DB 721 CTACAGAGAGCGGCGGAGGCGCTGGGGCGAGCGGCTGGCGGATGAGAGATGGGC 780
 QY 241 SerArgThrArgAspArgLeuAspGIuValLysGIuGIuValAlaGIuValArgAlaLys 260
 DB 781 AGCCGAGACCGCGGACCGCTGGAGAGGTGAAGAGAGAGGTGGCGGAGTGGCGCCCAAG 840
 QY 261 LeuGIuGIuGIuGIuGIuGIuGIuLeuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 280
 DB 841 CTGGAGGAGACGAGGCCACAGATACGCTGCAGGCGGAGGCTTCCAGGCCGCTCAAG 900
 QY 281 SerTrpPheGIuProLeuValGIuAspMetGIuArgGIuTrpAlaGIuLeuValGIuLys 300
 DB 901 ACCTGCTCGAGCCCTGCTGTGAGAGCATGACGCCAGTGGCGGCGGTGGTGAAGAAG 960
 QY 301 ValGIuAlaAlaValAlaGIuThrSerAlaAlaProValProSerAspAsnHis 317
 DB 961 GTGCAAGGCTGCGGTGGGACACGAGCGCGCCCTGTGCCACAGCAATCAC 1011
 RESULT 7
 AAD22048 ID AAD22048 standard; DNA; 1156 BP.
 AC AAD22048;
 XX
 DT 12-FEB-2002 (first entry)
 XX
 DE Human apolipoprotein E (apoE) isoprotein, apoE3 DNA.
 XX
 KW Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;
 XX hypertriglyceridaemia; low density lipoprotein; LDL; ds.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH CDS 61..1014
 FT /tag= a
 FT /product= "Human apoE isoprotein, apoE3"
 FT sig_peptide 61..114
 FT /tag= b

FT mat_peptide 115..1011
 FT /tag= c
 FT /product= "Mature human apoE isoprotein, apoE3"
 XX
 XX WO200177136-A1.
 XX
 XX 18-OCT-2001.
 XX
 XX 06-APR-2001; 2001WO-US11358.
 XX
 XX 06-APR-2000; 2000US-0544386.
 PR 04-OCT-2000; 2000US-0679088.
 PR 05-APR-2001; 2001US-0827854.
 XX
 XX (KOSP-) KOS PHARM INC.
 PA (UYBO-) UNIV BOSTON.
 PA
 XX Zannis VI, Kyreos KE;
 PI
 XX WPI: 2002-010885/01.
 DR P-PSDB; AAE13294.
 XX
 XX New apolipoprotein E polypeptide and nucleic acid, useful for lowering
 PT cholesterol, delaying the onset of or treating atherosclerosis in
 PT mammal, without inducing hypertriglyceridaemia
 XX
 XX Claim 14; Page 81; 91pp; English.
 XX
 XX The present sequence is a human apolipoprotein E (apoE)
 CC isoprotein, apoE3 DNA. The apoE lipoproteins are useful for
 CC lowering cholesterol, delaying the onset of atherosclerosis,
 CC treating or regressing atherosclerosis without inducing
 CC hypertriglyceridaemia, in a mammal lacking an endogenous,
 CC normally functioning apoE gene or low density lipoprotein (LDL)
 CC receptor or is at risk for developing atherosclerosis due to
 CC accumulation of lipoprotein remnants in the bloodstream or having
 CC a defect in remnant removal.
 XX
 SO Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other.
 Alignment Scores:
 Pred. No.: 1,9e-111 Length: 1156
 Score: 1573.00 Matches: 316
 Percent Similarity: 99.68% Conservative: 0
 Best local Similarity: 99.68% Mismatches: 1
 Query Match: 99.49% Indels: 0
 DB: 24 Gaps: 0
 US-09-827-854-14 (1-317) x AAD22048 (1-1156)
 QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGIuGIuAlaLysVal 20
 DB 61 ATGAAGGTTCTGTGGGCTGCGTGTGTCACATTCTGGCAGAGATGCCAAGGTG 120
 QY 21 GIuGIuAlaValAlaGIuThrGIuProGIuProGIuLeuArgGIuGIuTrpGIuSer 40
 DB 121 GAGCAAGCGGTGGAGACAGACCCGAGCGCGCCAGCCAGACAGACCGAATGGAGAGC 180
 QY 41 GIuGIuArgTrpGIuLeuAlaLeuGIuLysArgPheTrpAspTyrLeuArgTrpValGIuThr 60
 DB 181 GGCAGGCGCTGGAACTGGCACTGGTGTGCTTTGGGATTAAGTGGCTGGGTGAGACA 240
 QY 61 LeuSerGIuGIuValGIuGIuGIuLeuLeuSerGIuValThrGIuGIuLeuArgAla 80
 DB 241 CTGTCTGAGCAGGTGACAGAGAGCTGCTCAGTCCAGTCCAGAGAGAGAGAGGCG 300
 QY 81 LeuMetAspGIuTrpMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGIuLeu 100
 DB 301 CTGATGGACGACCATGAAAGAGTTGAAGGCTTCAAAATCGGAAGCTGAGAGCAACTG 360
 QY 101 ThrProValAlaGIuGIuThrArgAlaArgLeuSerLysGIuLeuGIuAlaGIuAla 120
 DB 361 ACCCGGTGGCGGAGAGACGCGGGCAGCGCTGTCCAAAGAGCTGAGGCGCGAGGCC 420

QY 121 ArgLeuGluGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyGluVal 140
 DB 421 CGGCTGGGCGCGGACATGAGGAGACGTGTGGCGGCTGTGTGACGTACCGCGGAGAGTG 480
 QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 481 CAGGCCATGCTGGCCAGACACCGAGAGACTCGGGTGGCTCCGCTCCGCTCCGCTCCG 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 DB 541 AAGCTGGGTAAAGGGCTCTCCCGGATGCCGATGACCTGAGAAAGCCCTGGGAGTGTAC 600
 QY 181 GlnAlaGlyAlaArgGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgGly 200
 DB 601 CAGCGCGGGCGCGGAGGGCGCGGAGCGGGCTCAGCGGCTCAGCCATCCGCGAGCGCTGGG 660
 QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 DB 661 CCCCTGGTGAAACAGGGCCCGCTGGCGGCGCCACTGTGGCTCCCTGGCGCGGCGAGCCG 720
 QY 221 LeuGlnGluArgAlaGlnAlaTyrPglGlyArgLeuArgAlaArgMetGluGlnMetGly 240
 DB 721 CTACAGAGAGCGGGCGCGGCTGGGCGAGCGGCTGGCGCGGATGGAGAGATGGGC 780
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
 DB 781 AGCGGAGCGCGGACCGCTGAGGAGGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
 QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 DB 841 CTGGAGAGAGAGCGCGGAGCATACGCTGACGCGGAGCGGCTCCAGGCGCGCTCCAG 900
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 DB 901 AGCTGGTTCGAGCCCTGGTGGAGAGACATCAGCGCCAGTGGGCGGGGCTGGTGGAGAG 960
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 961 GTGACAGGCTGCGTGGGACACGCGCGGCTGTGGCCAGGACATATC 1011
 RESULT 8
 AAD22052
 ID AAD22052 standard; DNA; 1156 BP.
 AC AAD22052;
 XX
 DT 12-FEB-2002 (first entry)
 XX
 DE Human apolipoprotein E (apoE) allele, apoE2** DNA.
 XX
 KW Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;
 KW hypertriglyceridaemia; low density lipoprotein; LDL; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 61..1014
 FT /tag= a
 FT /product= "Human apoE allele, apoE2**"
 FT sig_peptide 61..114
 FT /tag= b
 FT mat_peptide 115..1011
 FT /tag= c
 FT /product= "Mature human apoE allele, apoE2**"
 XX
 PN WO200177136-A1.
 XX
 PD 18-OCT-2001.
 XX
 PF 06-APR-2001: 2001WO-US11358.
 XX
 PR 06-APR-2000: 2000US-0544386.
 PR

PR 04-OCT-2000: 2000US-0679088.
 PR 05-APR-2001: 2001US-0827854.
 XX
 PA (KOSP-) KOS PHARM INC.
 PA (UYBO-) UNIV BOSTON.
 XX
 PI Zannis VI, Kypros KE;
 XX
 DR WPI: 2002-010885/01.
 DR P-PSDB: AAE13298.
 XX
 PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
 PT cholesterol, delaying the onset of or treating atherosclerosis in
 PT mammal, without inducing hypertriglyceridemia
 PS Claim 14; Page 83; 91pp; English.
 XX
 CC The present sequence is a human apolipoprotein E (apoE)
 CC allele, apoE2** DNA. The apoE lipoproteins are useful for
 CC lowering cholesterol, delaying the onset of atherosclerosis,
 CC treating or regressing atherosclerosis without inducing
 CC hypertriglyceridaemia, in a mammal lacking an endogenous,
 CC normally functioning apoE gene or low density lipoprotein (LDL)
 CC receptor or is at risk for developing atherosclerosis due to
 CC accumulation of lipoprotein remnants in the bloodstream or having
 CC a defect in remnant removal.
 XX
 SQ Sequence 1156 BP; 207 A; 369 C; 432 G; 148 T; 0 other:
 Alignment Scores:
 Pred. No.: 3,83e-111 Length: 1156
 Score: 1569.00 Matches: 315
 Percent Similarity: 99.68% Conservative: 1
 Best Local Similarity: 99.37% Mismatches: 1
 Query Match: 99.24% Indels: 0
 DB: 24 Gaps: 0
 US-09-827-854-14 (1-317) x AAD22052 (1-1156)
 QY 1 MetLysValLeuTrpAlaIleLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 61 ATGAAAGTTCTGTGGCTGTGGCTGTGTGTCACATTCTTGGCAGATGCCAGGCCAAAGTG 120
 QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 121 GAGCAAGCGGTGGAGAGACAGCGCGGAGCCGAGCTCGCGCAGAGACCGAGGCGAGAG 180
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB 181 GGCACAGCGCTGGGAACTGACAGGCGTCCCTTTGGATTACTCGCGTGGCGGACAGACA 240
 QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 DB 241 CTGTCTGAGCAGGTGAGAGAGAGGTGCTCAGCTCCAGGTACCCAGGAAGTGAAGGCG 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
 DB 301 CTGATGAGAGACCATGAAAGAGATTGAAGCGCTTCAAAATCGGAATGAGAGAACACTG 360
 QY 101 ThrProValAlaGlnGluThrArgAlaArgLeuSerLysGlnGlnAlaAlaGlnAla 120
 DB 361 ACCCGGTTGGCGAGAGAGAGCGCGGCGCTGTCCAAAGAGCTGAGCGCGCGAGGCC 420
 QY 121 ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyGluVal 140
 DB 421 CGGCTGGGCGCGGACATGAGGAGACGTGTGGCGGCTGTGTGACGTACCGCGGAGAGTG 480
 QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 481 CAGGCCATGCTGGCCAGACACCGAGAGACTCGGGTGGCTCCGCTCCGCTCCGCTCCG 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||||||:::|||||||||||||||||||||||||||||||||b|||||||

XX	RESULT 9	
XX	AAF84314	
XX	AAF84314 standard; cDNA; 1156 bp.	
AC	AAF84314;	
XX		
DT	21-JUN-2001 (first entry)	
XX		
DE	Human APOE2 coding sequence.	
XX		
KW	Human; APOE2; Alzheimer's disease; arteriosclerosis; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	61..1014
FT		/*tag= a
FT		/product= "Human APOE2"
XX		
FN	JP2001017028-A.	
XX		
PD	23-JAN-2001.	
XX		
PF	28-APR-2000; 2000JP-0128919.	
XX		
PR	06-MAY-1999; 99JP-0125647.	
XX		
PA	(MITU) MITSUBISHI CHEM CORP.	
XX		
DR	WPI: 2001-285406/30.	
DR	P-PSDB; AAB80996.	
XX		
PT	New APOE humanized mammalian cell useful for screening for agents	
PT	useful for treating or preventing Alzheimer's disease and	
PT	arteriosclerosis -	
XX		
PS	Disclosure; Page 11-12; 22pp; Japanese.	
XX		
CC	The present invention relates to an APOE humanised mammalian cell. The	
CC	present sequence is the coding sequence for human APOE2, which was used	
CC	in the method of the present invention. The APOE humanised mammalian cell	
CC	can be used for screening for agents useful for treating or preventing	
CC	Alzheimer's disease and arteriosclerosis.	

XX	Sequence	1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other:
Alignment Scores:		
Pred. No.:	7,75e-111	Length: 1156
Score:	1565.00	Matches: 315
Percent Similarity:	99.37%	Conservative: 0
Best Local Similarity:	99.37%	Mismatches: 2
Query Match:	98.9%	Indels: 0
DB:	22	Gaps: 0
US-09-827-854-14 (1-317) x AAF84314 (1-1156)		
QY	1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIleVal	20
DB	61 ATGAAGGTTCTGTGGGCTGCGCTGTGTGTCATTTCTCGGAGGATGCCAGGCCAAGGTG	120
QY	21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer	40
DB	121 GAGCAAGCGGTGGAGACAGAGAGCCGAGCCGAGCTGCCAGACAGACCGAGTGCAGAGC	180
QY	41 GlyGlnArgTTPGluLeuAlaLeuGluYarPheTrpAspTyrLeuArgTTPValGlnThr	60
DB	181 GGCACGACCTGGGAACCTGGCACTGGGTGCTTTGGATTTCCThCGCGTGGGTCCAGACA	240
QY	61 LeuSerGluGlnValGluGluLeuLeuSerSerGlnValTTrnGlnGluLeuArgAla	80
DB	241 CTGCTTGAGCAGAGTGCAGAGAGAGCTGCTCAGCTCCACAGTACCCAGAACTAGAGGCG	300
QY	81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu	100
DB	301 CTGATGGAGCAGACCATGAAGAGATTGAAGGCTTCAAAATCGAACTGAGAGCAACTG	360
QY	101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIleAlaGlnAla	120
DB	361 ACCCGGTCGGGAGGAGAACCGGGCACGGCTGTCCAGAGAGCTGACGGCGCCAGGCGC	420
QY	121 ArgLeuGluAlaAspMetGluAspValaArgGluYarGluLeuValGlnTyrArgGluVal	140
DB	421 CGCGTGGGCGCGGACATGAGAGAGCTGTGGGGCGCGCTGTGTGCACTACCGCGGCGAGGTG	480
QY	141 GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValaArgLeuAlaSerHisLeuArg	160
DB	481 CAGGCCATGCTTCGGCAAGACACGAGAGAGCTGGGGTGCCTTCGCTCCACCTCGC	540
QY	161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnTyrLysArgLeuAlaValTyr	180
DB	541 AAGCTGCGGTAAAGCGCTCTCCGCGATCCCATGATCCTCGAGAAAGTGCCTGCGAGTAC	600
QY	181 GlnAlaGluAlaArgGluGluYalAGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200
DB	601 CAGGCCGGGGCGCGGAGGGGCGGAGCGCGCTTCAGCGCCATCCCGGAGCGCTGGGG	660
QY	201 ProLeuValGluGlnGluYarValaArgAlaIleThrValaGlySerLeuAlaGluPro	220
DB	661 CCCCTGTGTGAAACAGGGCGCGTGGGGCCCCCCTGTGGGCTTCCGTGGCGGCCAGCG	720
QY	221 LeuGlnGluArgAlaGlnAlaIleTrpGluYalArgLeuArgAlaArgMetGluGluMetGly	240
DB	721 CTACAGGAGCGGGCCAGCGCTGGGGCGAGCGGCTGCGCGCGCGGAGATGAGATGGCC	780
QY	241 SerArgThrArgAspArgLeuAspGluValaLysGluGlnValaIleGluValaArgAlaLys	260
DB	781 AGCCGAGCCCGCGACCGCTGTGACAGGTGAAGAGCAGGTGGGGAGGTCCGGCCAG	840
QY	261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys	280
DB	841 CTGAGAGAGAGCGCCACAGATACGCTTCGAGGCCAGAGGCTTCCAGGCCCGGCTCAAG	900
QY	281 SerTrpPheGluProLeuValGluAspMetGlnArgGluTrpAlaGlyLeuValGluLys	300
DB	901 AGCTGGTTCGAGCCCTGTGTGAAGACACTGTGACAGCGCCAGTGGCGGGGCTGTGTGAGAG	960

XX Human apolipoprotein E (apoE) allele, apoE2* DNA.
DE Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;
KW hypertriglyceridaemia; low density lipoprotein; LDL; ds.
XX
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 61..1014
FT /*tag= a
FT /product= "Human apoE allele, apoE2*"
FT sig_peptide 61..114
FT /*tag= b
FT mat_peptide 115..1011
FT /*tag= c
FT /product= "Mature human apoE allele, apoE2*"
XX
XX W020177136-A1.
XX
XX 18-OCT-2001.
XX
XX 06-APR-2001; 2001WO-US11358.
XX
XX 06-APR-2000; 2000US-0544386.
XX 04-OCT-2000; 2000US-0679088.
XX 05-APR-2001; 2001US-0827854.
XX
XX (KOSP-) KOS PHARM INC.
XX (UYBO-) UNIV BOSTON.
XX
XX Zannis VI, Kyreos KE;
XX
XX WPI: 2002-010885/01.
XX P-PSDB: AAEL3297.
XX
XX New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT cholesterol, delaying the onset of or treating atherosclerosis in
PT mammal, without inducing hypertriglyceridaemia
XX
XX
XX Claim 14; Page 82; 91pp; English.
XX
XX The present sequence is a human apolipoprotein E (apoE)
CC allele, apoE2* DNA. The apoE lipoproteins are useful for
CC lowering cholesterol, delaying the onset of atherosclerosis.
CC treating or regressing atherosclerosis without inducing
CC hypertriglyceridaemia, in a mammal lacking an endogenous,
CC normally functioning apoE gene or low density lipoprotein (LDL)
CC receptor or is at risk for developing atherosclerosis due to
CC accumulation of lipoprotein remnants in the bloodstream or having
CC a defect in remnant removal.
XX
XX Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;
SQ

Alignment Scores:
Pred. No.: 7,75e-111 Length: 1156
Score: 1565.00 Matches: 315
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 98.99% Indels: 0
Gaps: 0

US-09-827-854-14 (1-317) x AAD22051 (1-1156)

QY 1 MetlySValLeuTRPAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGGTTCTGTGGGCGCTGCTGTCACATTCCGCGAGATGCCAGCAAGGTG 120
QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAAGCGGTGAGACGAGCCGAGCGCCGAGCGCCGAGCCGAGTGGCGAGAGC 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60

Db 181 GGCAGCGCTGGAGACTGCGACGTGGTGGCTTTGGGATFACCTGGCTGGGTGAGACA 240
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGTCTGAGCAGGTGACAGAGAGAGCTGCTCAGCTCCAGCTCCAGCAGAACTGAGGGCG 300
QY 81 LeuMetAspGluTrpMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
Db 301 CTGATGACGAGACCATGAAGAGGTGAGAGCCCTTCAATATGGAAGTGGAGAACACTG 360
QY 101 ThrProValAlaGluGluTrpThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 361 ACCCGGTGGCGAGAGACCGCGGCGACGCTGTCCAGAGAGCTGACAGCGCGCGAGCC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyVal 140
Db 421 CGGCTGGGCGCGACATGAGAGAGCTGTGCGCGCCCTGGTGCATACCGCGGAGGTG 480
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGCGCATGCTCGGCGCAGAGACCGAGAGCTCGCGGTGCGCTTCCACCTGGCGC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTGTGTAAGCGGCTCTCCGCGATGCCATGACCTGCAGAAAGCGCTGGCACTGAC 600
QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 601 CAGCGCGGCGCGCGAGGCGCGAGCGCGCGCTTCCAGCTCCGCGAGCGCTTGGCG 660
QY 201 ProLeuValGluGlnGlyArgValArgValArgAlaThrValGlySerLeuValGlyGlnPro 220
Db 661 CCCCTGGTGAACAGGCGCGCGGTGGGCGCCACTGTGGCTTCCCTGGCGGCGAGCGG 720
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
Db 721 CTACAGAGAGCGGCGCCAGGCTCGGGGCGAGCGGCTGCGCGCGGATGGAGAGATGGCG 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 781 AGCGGAGCCCGACCGCTGCGAGGTGAAGAGCGAGGTGGGAGGTGCGCGCCAG 840
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgPheGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 841 CTGAGAGAGCGCGCCAGCATACGCTGCGAGCGCGAGGCTTCCAGGCCGCTCAAG 900
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 901 AGCTGTTCGAGCCCTGCTGTGAAGACATGACGCGCCAGTGGCGCGGCTGTGGAGAAAG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCGTGGGACACCGCGCCCTGTGCTCCAGCGACGAATCAC 1011

RESULT 12
AAT06957
ID AAT06957 standard; cDNA to mRNA; 1157 BP.
XX
AC AAT06957;
XX
DT 19-JUN-1996 (first entry)
XX
XX Human apolipoprotein-E (ApoE) cDNA.
DE
XX Recombinant; human; apolipoprotein-E; ApoE; insect cells; larva;
KW Manduca sexta; Autographa californica nuclear polyhedrosis virus;
KW haemolymph; lipid complex; biologically active; ds.
XX
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc_feature 62..119


```

FT      /tag= a
FT      note= "misc_signal"
FT      CDS      62..1015
FT      /tag= b
XX      US5472858-A.
XX      PD      05-DEC-1995.
XX      PF      04-JUN-1991: 91US-0709949.
XX      PR      04-JUN-1991: 91US-0709949.
XX      PA      (MISC ) WISCONSIN ALUMNI RES FOUND.
XX      PI      Attie AD, Beckage NE, Gretch DG, Sturley SL;
XX      DR      WPI; 1996-029812/03.
XX      DR      P-PSDB; AAR86791.
XX      PT      Prodn. of recombinant apolipoprotein E in insects - by infecting
PT      Manduca sexta larvae with recombinant Autographa californica
PT      nuclear polyhedrosis baculovirus vector.
XX      PS      Disclosure; Columns 11-14; 10pp; English.
XX      CC      Recombinant human apolipoprotein-E (APOE) (AAR86791) can be produced
CC      by preparing a genetic construct (contg. an APOE-encoding sequence,
CC      e.g. AAT06957, and flanking regulatory sequences enabling the protein
CC      to be expressed in insect cells), which is then introduced into a
CC      Manduca sexta larva (using a recombinant Autographa californica
CC      haemolymph of the larval host. The APOE produced is in a form
CC      sufficiently complexed with lipids to be biologically active, which
CC      cannot be achieved in insect cell cultures, and can therefore be
CC      used in therapeutic applications.
XX      SQ      Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;

Alignment Scores:
Pred. No.: 9.24e-111 Length: 1157
Score: 1564.00 Matches: 314
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 99.05% Mismatches: 3
Query Match: 98.92% Indels: 0
DB: 17 Gaps: 0

US-09-827-854-14 (1-317) x AAT06957 (1-1157)
QY      1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaTysVal 20
DB      62 ATGAAGATTCTCTGGGCTGGTGGTGGTGCACATTCTCTGGCAGAGATGCCAGGCGCAAGTG 121
QY      21 GlnGlnAlaValAlaGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnThrGlnSer 40
DB      122 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTCGGCGAGAGCCAGTGGCGAGAC 181
QY      41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB      182 GGCAGCGCTGGGAACCTGGCAGCTGGTCTTTGGGATTACCTGGCGGTGGCGAGATA 241
QY      61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB      242 CTGTCTGACGACAGTGCAGAGAGAGTGTCTCAGCTCCCAAGTCCCAAGAACTGAGAGGCG 301
QY      81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
DB      302 CTGATGGAGAGACCACTGAAGAGGTGAAGGCTTACAAATCGGAATCGAGGAACTG 361
QY      101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
DB      362 ACCCCGCTAGCGGAGAGAGAGCGGCGCTGTCCAAAGGAGCTGCAGACGGCGAGGCC 421

```

```

QY      121 ArgLeuGlyAlaAspMetGlnAspValArgGlyArgLeuValGlnThrArgGlyGlnVal 140
DB      422 CGGCTGGCGCGGAGACATGAGAGAGCTGTGGGCGCGCTGGTGTGACGACCGCGAGTG 481
QY      141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB      482 CAGGCCATGCTGGCCAGAGACCGAGAGCTTCGGGTGGCGCTCCCTCCCACTGCGCC 541
QY      161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB      542 AACCTGCTGTAAGGAGCTCTCTCCGATCCGATGACCTGCAGAGAGCGCTGGCAGTGTAC 601
QY      181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB      602 CAGCGCGGGGCGCGGAGGCGCGGAGCGCGGCGCTCAGCGCCATCCGAGCGCGCTGGGG 661
QY      201 ProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
DB      662 CCCCTGTGTGAAACAGCGCGCGCTGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
QY      221 LeuGlnGlnAlaArgAlaGlnAlaTrpGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 240
DB      722 CTACAGAGAGGAGCCCAAGGCTGGGAGCGGCTGCGCGCGGAGAGAGAGATGGCC 781
QY      241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValAlaArgAlaLys 260
DB      782 AGTCGAGACCGCGACCGCTGGACGAGGTGAAGAGACAGTGGCGGAGGTGGCGCGCAAG 841
QY      261 LeuGlnGlnGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
DB      842 CTGAGAGACGAGCCCGACAGATATCCCTGACAGGCGCGAGCCCTTCAGGCGCGCTCAAG 901
QY      281 SerTrpPheGlnProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
DB      902 AGCTGTTTGCAGACCCCTGTGTGAAGACATGCAGCGCCAGTGGCGCGCTGTGTGAGAAAG 961
QY      301 ValGlnAlaAlaValGlnThrSerAlaAlaProValProSerAspAsnHis 317
DB      962 GTGCAGGCTGCGTGGGACACAGCGCGCGCTGTGTCACGACACATATCAC 1012

RESULT 13
ABN95746
ID      ABN95746 standard; DNA; 1157 BP.
AC      ABN95746;
XX      13-AUG-2002 (first entry)
DT      13-AUG-2002 (first entry)
DE      Gene #2244 used to diagnose liver cancer.
XX      Gene: liver cancer; ds; hepatocellular carcinoma; hepatotropic;
XX      metastatic liver tumour; cytostatic; expression profile; disease state;
XX      disease progression; drug toxicity; drug efficacy; drug metabolism.
OS      Homo sapiens.
PN      WO200229103-A2.
PD      11-APR-2002.
XX      02-OCT-2001; 2001WO-US30589.
XX      02-OCT-2000; 2000US-237054P.
XX      (GENE-) GENE LOGIC INC.
PA      Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
PI      WPI; 2002-426119/45.
XX      Diagnosing and detecting the progression of liver cancer.
PT      hepatocellular carcinoma or metastatic liver tumor in a patient.
PT      involves detecting the level of expression of two or more genes in a

```


PT liver tissue sample -
 XX
 PS Claim 1: SEQ ID NO 2244; 298bp; English.
 CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumour in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytosolic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;
 Alignment Scores:
 Pred. No.: 9 24e-111 Length: 1157
 Score: 1564.00 Matches: 314
 Percent Similarity: 99.05% Conservative: 0
 Best Local Similarity: 99.05% Mismatches: 3
 Query Match: 98.92% Indels: 0
 DB: 24 Gaps: 0
 US-09-827-854-14 (1-317) x ABN95746 (1-1157)
 QY 1 MettysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaAlaVal 20
 Db 62 ATGAAGGTTCTGTGGCGTGGTGTGTCACATTCTCGGAGAGATGGCAGCCAGGTG 121
 QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 Db 122 GAGCAAGCGGTGGAGACAGACCGGAGCCGAGCTGCCAGACAGACCGAGTGGCAGAC 181
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 Db 182 GGCACGCGCTGGAGACTGGCTGGTGGCTTTGGGATTCACCTGGCTGGGTGCAGACA 241
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 Db 242 CTGCTGTGACGAGTGCAGAGAGAGCTGCTCAAGTCCCAAGTCCACCAAGAACTGAGGCG 301
 QY 81 LeuMetAspGluThrMetLeuGluLeuLysAlaTrpTrpLysSerGluLeuGluGlnLeu 100
 Db 302 CTGATGGACGAGACCATGAGAGAGTTGAAGGCTTACAAATCGGAACCTGGAGAACAACTG 361
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
 Db 362 ACCCGGTAGGAGAGACACCGGCGAGCTGTCCAAAGAGCTGCACACGCGCCAGGCC 421
 QY 121 ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTrpArgGluVal 140
 Db 422 CGGCTGGGCGCGGACATGAGAGAGCTGTGGCGCCCGCTGTGCTGCTACCGCGGAGAGTGTG 481
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 482 CAGGCAATGCTCGGCAAGACACCGAGAGCTGGGGGCGCCCTCCGCTCCACCTGGCGC 541
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValLys 180
 Db 542 AAGCTGCTAAGCGGCTCTCCGCGATCCGATACCTGCAAGCGGCTGCGAGTGTAC 601
 QY 181 GlnAlaGlyAlaArgGluGluValAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db 602 CAGGCGGGGCGCGGAGAGGCGCGGAGCGGCTCAGCGCCATCCGAGAGCGCTGGGG 661
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220

Db 662 CCCCTGGTGAACAGGCGCGGTGGCGCCGCACTGTGGGCTCTGGCGCGCAGCCG 721
 QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
 Db 722 CTACAGGAGCGGGCCAGGCTTGGGGGAGCGGCTGGCGCGGATGTGAGAGATGGGC 781
 QY 241 SerArgTrpArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 Db 782 AGTGGAGACCGCGACCGCTGTGAGAGTGAAGAGACAGGAGTGGGAGAGTGGCGCCAG 841
 QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 Db 842 CTGGAGGAGCGAGCGCCACAGATACGCTTCAGCGCGGCGGCTTCAGAGCCGCTCAAG 901
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 Db 902 AGCTGGTTCAGGCCCTCTGTGAAAGACATGCAGGCGCAGTGGGCGGCTGTGGAGAG 961
 QY 301 ValGlnAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 962 GTGCAGGCTGCGGTGGGACAGCGCGCCCTGTGGCCAGCAGCAATCAG 1012
 RESULT 14
 ABK64514
 ID ABK64514 standard; DNA; 1157 BP.
 AC ABK64514:
 XX 18-JUN-2002 (first entry)
 DT Human benign prostatic hyperplasia gene #409.
 DE Human: benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
 KW Homo sapiens.
 OS WO200212440-A2.
 PN 14-FEB-2002.
 PD 07-AUG-2001; 2001WO-US24708.
 PE 07-AUG-2000; 2000US-223323P.
 PR 05-JUN-2001; 2001US-0873319.
 XX (GENE-) GENE LOGIC INC.
 PA (MISB) JAPAN TOBACCO INC.
 PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
 XX WPL; 2002-257476/30.
 DR Identifying drugs for and diagnosing benign prostatic hyperplasia, by
 PT detecting expression levels of one or more genes in prostate cells from
 PT patient that are differentially regulated compared to normal prostate
 PT cells -
 PS Disclosure; page 239-240; 444pp; English.
 XX
 CC The invention relates to a method of diagnosing (I) the onset or
 CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
 CC or identifying an agent that modulates the onset or progression of BPH.
 CC The method is based on changes in gene expression in BPH tissue isolated
 CC from patients exhibiting different clinical states of prostate
 CC hyperplasia as compared to normal prostate tissue. (I) comprises
 CC detecting the expression levels of one or more genes in prostate cells
 CC from the subject that are differentially regulated compared to normal
 CC prostate cells. (II) comprises preparing a first gene expression profile
 CC of BPH cells or BPH-like cell population, exposing the cells to the
 CC agent, preparing a second gene expression profile of the agent exposed
 CC cells, and comparing the first and second gene expression profiles.
 CC (I) is useful for diagnosing the onset or progression of BPH. (II) is

PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppel DR, Weaver Z;
 DR WPI: 2002-188264/24.
 XX
 XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 Claim 1: SEQ ID 3787; 44pp; English.

XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 CC
 XX

XX Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;

Alignment Scores:

Pred. No.: 9,24e-111 Length: 1157
 Score: 1564.00 Matches: 314
 Percent Similarity: 99.05% Conservative: 0
 Best Local Similarity: 99.05% Mismatches: 3
 Query Match: 98.92% Indels: 0
 DB: 24 Gaps: 0

US-09-827-854-14 (1-317) x ABL65450 (1-1157)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 62 ATGAAGGTTCTGTGGCTGCTTCTGTCACTTCTGGCAGAGATGCCAAGCCAAAGTGT 121
 QY 21 GlnGlnAlaValGlnTrpGlnProGlnProGlnLeuArgGlnGlnTrpGlnSer 40
 DB 122 GAGCAACGCGTGGAGACAGACCGGAGCTGGCCAGCAGACCGAGTGGCAGACG 181
 QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB 182 GGCACGCGGTGGGAACGTCGCTGCTTTGGGATTACCTGGCTGGTGCACAGCA 241
 QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 DB 242 CTGCTCTAGCAGAGTGGCAGGAGAGCTCTCACTCCCAAGTCAACCAAGATGAGGGCG 301
 QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
 DB 302 CTGATGACGACGACCATGAGAGACTTGAAGCCCTACAAATCGGAATCGAAGCAACTG 361
 QY 101 ThrProValAlaGlnGlnGlnTrpArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
 DB 362 ACCCGGTAGCGAGGAGACGCGGCGACGGCTGTCCAAAGAGATGACAGACGCGCAGGCC 421

QY 121 ArgLeuGlnAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyGluVal 140
 DB 422 CGCGTGGGCGCCGACATGAGAGACGTGTGGGCGCCGCTGTGACATGCCGGCGAGATG 481
 QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 482 CAGGCCATGCTCGGCCAGAGACACCGAGAGAGTGGGGTGGCCCTCGCCCTCCACCTCGCG 541
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 DB 542 AAGCTGGCTAAGCGGCTCTCCGCGATCCGATACCTGCAAGACCGCTGGCGAGTGTAC 601
 QY 181 GlnAlaGlnAlaArgGlnGlnGlnAlaGlnArgGlyLeuSerAlaLeuArgGlnArgLeuGly 200
 DB 602 CAGGCGCGGCGCCCGAGGCGCCGAGCGGCTCAAGCCCATCCCGAGAGCGCTGGG 661
 QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlnSerLeuAlaGlnGlnPro 220
 DB 662 CCCCTGGTGACAGGCGCGCTGGCGGCGCCACTGTGGGCTCCCTGGCGGCGCAGCCG 721
 QY 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
 DB 722 CTACAGGAGGCGGCGCCAGGCTGGGCGAGCGCTGGCGCGCGGATGAGAGATGGCG 781
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaLys 260
 DB 782 AGTGGACCCGCGACCCCTGGAGAGGTGAAGAGAGGTGGCGGAGTGGCGGCTCAAG 841
 QY 261 LeuGlnGlnGlnAlaGlnGlnGlnLeuArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 DB 842 CTGAGAGAGCAGGCGCCAGCAGATACGCTCAGGCGAGGCTTCCAGGCGCGGCTCAAG 901
 QY 281 SerTrpPheGlnProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
 DB 902 AGCTGGTTCCAGGCCCTCGGTGGGAAGACATGCAGGCGCAGTGGCGGCTGGTGGAGAG 961
 QY 301 ValGlnAlaAlaValGlnGlnThrSerAlaAlaProValProSerAspAsnHis 317
 DB 962 GTGCAGGCTGCGCTGGGACACAGCCGCCCTGTGCTGCGCCAGCAGCAATTCAC 1012

Search completed: March 14, 2003, 12:42:52
 Job time : 184.728 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:17:52 : Search time 34.7343 Seconds
(without alignments)
2798.866 Million cell updates/sec

Title: US-09-827-854-14

Perfect score: 1581
Sequence: 1 MKVLMALVLEFLAGCQAKV.....VEKQAAVGVTSAPVPSDNH 317

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_plus.p2n.model -DEV=xlp
-O=/cgn2.1/USPTO/spool/US09827854/runat_11032003_101611_27510/app.query.fasta.1.3576
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptio -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09827854.ecgn1.1.1224.runat_11032003_101611_27510 -NCPU=6 -ICPU=3
-NO_XLPTX -NO_MMAP -LARGEQUERY -NEGSCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120
-WARN-TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Issued_Patents_NA:*
2: /cgn2.6/ptodata/2/ina/5A.COMB.seq:*
3: /cgn2.6/ptodata/2/ina/5B.COMB.seq:*
4: /cgn2.6/ptodata/2/ina/6A.COMB.seq:*
5: /cgn2.6/ptodata/2/ina/6B.COMB.seq:*
6: /cgn2.6/ptodata/2/ina/PCRTUS.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1564	98.9	1157	1	US-07-709-949-1
2	1146	72.5	1126	4	US-08-949-155-5
3	1146	72.5	1126	4	US-09-819-964-5
4	1010	63.9	4267	4	US-08-949-155-51
5	1010	63.9	4267	4	US-09-819-964-51
6	912	61.5	660	2	US-08-726-306A-28
7	519	32.8	330	1	US-07-849-389-6
8	363	23.0	252	3	US-08-617-256-24
9	363	23.0	252	4	US-09-287-141-24
10	363	23.0	252	4	US-09-431-613-24
11	363	23.0	252	4	US-09-504-245-24
12	363	23.0	252	4	US-09-287-682-24

13	363	23.0	252	4	US-09-287-679-24	Sequence 24, Appl
14	363	23.0	252	4	US-09-397-766-24	Sequence 24, Appl
15	363	23.0	252	4	US-09-287-681-24	Sequence 24, Appl
16	363	23.0	252	4	US-09-495-444-24	Sequence 24, Appl
17	182.5	11.5	842	4	US-08-952-796-1	Sequence 4, Appl
18	179.5	11.4	801	1	US-07-959-946-4	Sequence 1, Appl
19	179.5	11.4	801	1	US-08-333-577-4	Sequence 4, Appl
20	179.5	11.4	801	5	PCT-US92-08634-4	Sequence 4, Appl
21	162.5	10.3	964	1	US-08-448-606-5	Sequence 5, Appl
22	159.5	10.1	863	1	US-08-448-606-7	Sequence 7, Appl
23	158	10.0	1879	4	US-09-750-580-2	Sequence 2, Appl
24	153	9.7	3256	2	US-08-968-751-3	Sequence 3, Appl
25	152.5	9.6	603	4	US-08-952-796-14	Sequence 14, Appl
26	150.5	9.5	1771	2	US-08-533-669A-7	Sequence 7, Appl
27	150.5	9.5	1771	2	US-08-511-872-1	Sequence 1, Appl
28	150.5	9.5	1771	4	US-09-183-861-7	Sequence 7, Appl
29	150.5	9.5	1771	4	US-09-022-765-7	Sequence 7, Appl
30	149	9.4	516	4	US-09-183-861-34	Sequence 34, Appl
31	149	9.4	516	4	US-09-022-765-34	Sequence 34, Appl
32	149	9.4	6306	1	US-08-195-487-3	Sequence 3, Appl
33	149	9.4	6306	5	PCT-US93-06160-3	Sequence 3, Appl
34	149	9.4	8789	1	US-08-328-234-5	Sequence 5, Appl
35	149	9.4	10136	5	PCT-US95-16216-2	Sequence 2, Appl
36	149	9.4	10136	5	PCT-US95-16216-2	Sequence 12, Appl
37	148	9.4	4868	1	US-08-139-937-12	Sequence 12, Appl
38	148	9.4	4868	5	PCT-US93-11310-12	Sequence 12, Appl
39	148	9.4	13121	4	US-08-961-527-126	Sequence 126, Appl
40	147	9.3	5661	4	US-08-938-105-2	Sequence 2, Appl
41	145	9.2	1960	2	US-08-533-306A-1	Sequence 1, Appl
42	145	9.2	1960	2	US-08-742-923A-1	Sequence 1, Appl
43	145	9.2	6306	1	US-08-466-390-3	Sequence 3, Appl
44	145	9.2	6306	1	US-08-470-950-3	Sequence 3, Appl
45	145	9.2	6306	1	US-08-467-781-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-07-709-949-1
Sequence 1, Application US/07709949
Patent No. 5472858
GENERAL INFORMATION:
APPLICANT: Attie, Alan D
APPLICANT: Gretch, Daniel G
APPLICANT: Sturley, Stephen L
APPLICANT: Beckage, Nancy E
TITLE OF INVENTION: Production of Recombinant Proteins in
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: P.O. Box 2113
CITY: Madison
STATE: WI
COUNTRY: USA
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/709,949
FILING DATE: 19910604
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 9629691801
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 1:


```

NAME/KEY: CDS
LOCATION: 51..1001
US-08-949-155-5

Alignment Scores:
Pred. No.: 4,98e-101
Score: 1146.00
Percent Similarity: 84.06%
Best Local Similarity: 70.31%
Query Match: 72.49%
DB: 4
Gaps: 3

US-09-827-854-14 (1-317) x US-08-949-155-5 (1-1126)

OY 1 MetcysValLeuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 51 ATGAGGGTTCTGTGGCTTGTGGTACCTCTCTCGCAGATGCGGACAGAGAC 110
OY 21 GluGlnAlaValAluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
DB 111 GAGCGCGGG-----CCGCGCGCGAGGTGCACGTGTGTGGAGAGAGCCCAAG 158
OY 38 TrpGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrp 57
DB 159 TGGCAGGCGCAGCCAGCCCTGGAGCAGCCCTGGCGCTCTGGGATTACCTGGCGCTGG 218
OY 58 ValGlnThrLeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGlu 77
DB 219 GTGCGAGTCCCTGTCTACCAAGTGCAGAGGAGAGCTCTAGACCAAGGTCAAGGAGAA 278
OY 78 LeuArgAlaLeuMetAspLysThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlu 97
DB 279 CTGACGCGACTGATAGAGAGAGACATGAGAGAGTGAAGCCTACCGCAGAGAGCTGGAG 338
OY 98 GluGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAla 117
DB 339 GCGCAGCTGGCCCTGACCCAGCAGCAGCAGCGCCGCTGCCAAGAGAGCTGCAGCGCG 398
OY 118 AlaGlnAlaArgLeuGlyAlaAspMetGluAspValArgGlyLeuSerAlaLeuArgGlu 137
DB 399 GCGCAGGCGCGGTGGGCGCCGACATGAGAGAGCTGCGACCCGCTGTGCTCTACCCG 458
OY 138 GlyGluValGlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSer 157
DB 459 ACGAGAGTGCACACATGTTGGGCGCAGCAGCAGAGAGACTCGGAGCGCCCTGGCTTCC 518
OY 158 HisLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeu 177
DB 519 CACCTGCGCAAGCTGCGCAAGCGGCTGCTCCGACACCGAGAGAGACTGCAGAAAGCGCTG 578
OY 178 AlaValTyrGlnAlaGlyAlaArgGluGlyAlaArgGlyLeuSerAlaLeuArgGlu 197
DB 579 GCGGTGTACAGAGCGGGGTGGCGCGAGGCGCGAGCGGTGAGCCCTCCGCGAG 638
OY 198 ArgLeuGlyProLeuValGluGlnGlyArgValArgAlaAlaThrArgLysSerLeuAla 217
DB 639 CCGCTGGGCGCCCTGTGTGAGAGAGAGCGCATTCGCGCGCCGACCTGTGATGCCAGGCGC 698
OY 218 GlyGlnProLeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGlu 237
DB 699 GCGCAGCGCCCTGCGGAGCGCGGAAAGCTGGGCGCAGAACCTGCGCGAGCGGCTGGAG 758
OY 238 GluMetGlySerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluVal 257
DB 759 GAGATGGGAGCGCGGACCGCGACCGCTGATGATGATCGTGAACAGTGAAGAGGAGTG 818
OY 258 ArgAlaLysLeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 277
DB 819 CCGACCAAGTGTGAGAGAGAGAGGAGCGACGTCGCGCTCAGAGCCGAGGAGATTCCAGCGC 878
OY 278 ArgLeuLysSerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGluLeu 297
DB 879 CTCCTCAAGGCTGCTGTGAGCGCTGCTGTGAGAAACATACGCGCGCAGTGGCGCGGCTG 938

```

```

OY 298 ValGluLysValGlnAlaValAlaGly---ThrSerAlaAlaProValProSerAspAsn 316
DB 939 GTGAGAGAGATCAGTGGCGGTGAGAGATTAACCTCTCCACTGTGCGCCAGATTAAT 998

RESULT 3
US-09-819-964-5
Sequence 5, Application US/09819964
Patent No. 6369294
GENERAL INFORMATION:
APPLICANT: Pledrahlta, Jorge A
Bazer, Fuller W
TITLE OF INVENTION: Compositions and Methods for the
Generation of Transgenic Animal Species
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESS: ARNOLD, WHITE AND DURKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/819,964
FILING DATE: 28-Mar-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/949,155
FILING DATE: <unknown>
APPLICATION NUMBER: US 60/046,094
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hibler, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TANK:177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 51..1001
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-819-964-5

Alignment Scores:
Pred. No.: 4,98e-101
Score: 1146.00
Percent Similarity: 84.06%
Best Local Similarity: 70.31%
Query Match: 72.49%
DB: 4
Gaps: 3

US-09-827-854-14 (1-317) x US-09-819-964-5 (1-1126)

OY 1 MetcysValLeuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 51 ATGAGGGTTCTGTGGCTTGTGGTACCTCTCTCGCAGATGCGGACAGAGAC 110
OY 21 GluGlnAlaValAluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
DB 111 GAGCGCGGG-----CCGCGCGCGAGGTGCACGTGTGTGGAGAGAGCCCAAG 158

```



```

OY 38 TTPGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTyr 57
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 159 TGGCAGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 218
OY 58 ValGlnThrLeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGlu 77
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 219 GTGCAAGTCCCTGTGACCAAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 278
OY 78 LeuArgAlaLeuMetAspGluThrMetGlyGluLeuValAlaTyrGlySerSerGluLeuGlu 97
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 279 CTGACGAGAGCTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 338
OY 98 GluGlnLeuThrProValAlaGlnGluThrArgAlaArgLeuSerGlyGluLeuAla 117
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 339 GCCCAGCTGGGCGCCGTGACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 398
OY 118 AlaGlnAlaArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArg 137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 399 GCGCAGGCGCGCGTGGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 458
OY 138 GlyGluValGlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSer 157
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 459 AGCGAGGTGCACAACATGTTGGGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 518
OY 158 HisLeuArgGlyLeuArgGlyArgLeuLeuArgAspAlaAspAspLeuGlnValArgLeu 177
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 519 CACCTGCGCAAGCTGCGCAAGCGCGTGCCTCCGAGACACCGAGAGAGAGAGAGAGAGAG 578
OY 178 AlaValTyrGlnAlaGlyAlaArgGluGlnGluAlaGlnArgGlyLeuSerAlaIleArgGlu 197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 579 GCCGTGTACCAAGCGCGGCGCTCCGCGAGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAG 638
OY 198 ArgLeuGlyProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAla 217
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 639 CGCGCTGGGCGCCCTGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 698
OY 218 GlyGlnProLeuGlnGlnArgAlaGlnAlaThrGlyGlnArgLeuArgAlaArgMetGlu 237
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 699 GCGCAGCGCGCTGGCGAGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 758
OY 238 GluMetGlySerArgThrArgAspArgLeuAspGluValGlyGlnGluValAlaGlnVal 257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 759 GAGATGGCGAGCGCGAGCGCGAGCGCGCTGGATGAGATGCGTGAAGCGAGAGAGAG 818
OY 258 ArgAlaLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 277
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 819 CGCACAAGAGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 878
OY 278 ArgLeuLeuSerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeu 297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 879 CTGCTCAAAAGGCTGTTCGAGCGCTGTGGTGAAGACATACGCGCGAGTGGCGCGGCTG 938
OY 298 ValGlnThrLeuSerGluGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 939 GTGGAAGAGATGACGTGCGCGGTGAGACATACGCTCTCCACTCTGCGCGCCAGTATAT 998

```

US-08-949-155-51
 : Sequence 51, Application us/08949155
 : Patent No. 6271436

GENERAL INFORMATION:

APPLICANT: Pledralita, Jorge A
 TITLE OF INVENTION: Compositions and Methods for the
 : TITLE OF INVENTION: Generation of Transgenic Animal Species
 : NUMBER OF SEQUENCES: 51
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: ARNOLD, WHITE AND DURKEE
 : STREET: P. O. Box 4433
 : CITY: Houston
 : STATE: TX
 : COUNTRY: US
 : ZIP: 77210-4433

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/949,155
: FILING DATE: Concurrently Herewith
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/027,338
: FILING DATE: 11-OCT-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/046,094
: FILING DATE: 09-MAY-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Hibler, David W.
: REGISTRATION NUMBER: 41,071
: REFERENCE/DOCKET NUMBER: TANK:177
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (512) 418-3000
: TELEFAX: (713) 789-2679
: INFORMATION FOR SEQ. ID NO: 51:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4267 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-949-155-51

```

```

Alignment Scores:
Pred. No.: 3.14e-87 Length: 4267
Score: 1010.00 Matches: 219
Percent Similarity: 58.65% Conservative: 42
Best Local Similarity: 49.21% Mismatches: 45
Query Match: 63.88% Indels: 140
DB: 4 Gaps: 5

```

US-09-827-854-14 (1-317) x US-08-949-155-51 (1-4267)

```

OY 1 MetGluValLeuTrpAlaAlaLeuValAlaThrPheLeuAlaGlyCysGlnAlaVal 20
    ||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2448 ATGCGTGTGACTGTGCTCGCT-----GCAGATGCGCGAGAGAGAGAGAGAG 2489
OY 21 GluGlnAlaValGlnThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
    ||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2490 GACCGCGGG-----CCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2537
OY 38 TTPGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTyr 57
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2538 TGGCAGGCGCAGCAGCGCTGGGAGCAGCGCGCTGGCGCTTCTGGGATTAACCTGCGCTGG 2597
OY 58 ValGlnThrLeuSerGluGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 77
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2598 GTGCAAGTCCCTGTGACCAAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2657
OY 78 Leu----- 78
Db 2658 CT-GACGTAAGTCCACACCGAGACTCCCGCGCGCGCGCGCGCGCGCGCGCTGA 2716
OY 78 ----- 78
Db 2717 CCCTCTGGGGAACCGTGTGTTCTGAGACCTCAGGCTCAGCCGTCGGGTTCTTCTG 2776
OY 78 ----- 78
Db 2777 TCCTTGTCCCAACTCTTGGGGGTCTGGGTCTGTCTTTCTTTTCTCTCTCTCTTT 2836
OY 78 ----- 78
Db 2837 TTGGGGGAAAAAATTTTCTTTTCTTTTCTTTCAATTGACTCATGCTTGTCTTCTCC 2896
OY 78 ----- 78

```



```

Db 2837 TTGGGGGAAAAAATTTTCTTTTCTTTCTTCAATTGACTTCAATGCTTCTTTCTTTCC 2896
Oy 78 ----- 78
Db 2897 ATCTGAGCTCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2956
Oy 78 ----- 78
Db 2957 AATCTGTGACGTCTGCTGCGCATGCGCAGTCAGAGCCCTCTCTCCCTCACCGCC 3016
Oy 79 ----- 79
Db 3017 CCCGCCCTCTCTGCGCCCGAGAGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3076
Oy 93 LysSerGluLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 112
Db 3077 CGCGAGAGAGCTGAGCGAGCGAGCTGCGGCCCGCTGACCCAGAGAGAGAGAGAGAGAGAG 3136
Oy 113 LysGluLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 132
Db 3137 AAGAGAGCTGAGCGCGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3196
Oy 133 LeuValGluThrArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 152
Db 3197 TTGCTCTCTACCGAGCGAGGCTGACAAACATTTGGGCGCAGACACCGAGAGCTGCGG 3256
Oy 153 ValArgLeuAlaSerHisLeuArgLysLeuArgLysLeuArgLysLeuArgLysLeuArg 172
Db 3257 AGCGGCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3316
Oy 173 LeuGluLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGluGluGluGluGluGlu 192
Db 3317 CTGCAAGAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3376
Oy 193 SerAlaIleArgGluArgGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 212
Db 3377 AGCGGCTGCGCGAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3436
Oy 213 ValGlySerLeuAlaGlyLeuProLeuGluGluGluGluGluGluGluGluGluGluGlu 232
Db 3437 CTGAGTACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3496
Oy 233 ArgAlaArgMetGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 252
Db 3497 CGGAGAGCGCTGAGAGAGATGGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3556
Oy 253 GlnValAlaGluValArgAlaLysLeuGluGluGluGluGluGluGluGluGluGluGlu 272
Db 3557 CAGCTGAGAGAGAGTGGCGCACCAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3616
Oy 273 GluAlaPheGlnAlaArgLeuLysSerTrpPheGluProLeuValGluAspMetGlnArg 292
Db 3617 GAGGATTCACCGCTCTCAAGGCTGTTGAGCTCTGTTGAGAGAGAGAGAGAGAGAGAG 3676
Oy 293 GlnTTPAlaGlyLeuValGluLysValGlnAlaAlaValGly---ThrSerAlaAlaPro 311
Db 3677 CAGTGGCGCGGCTGGTGGAGAGATGATGCGCGCTGAGATGATGCTCTCCACCTCT 3736
Oy 312 ValProSerAspAsn 316
Db 3737 GCGCCCGATGATAT 3751

```

```

ADDRESSER: Banner & Wilcoff, Ltd.
STREET: 1 Financial Center
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,306A
FILING DATE: 02-Oct-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 95/20080.4
FILING DATE: 02-Oct-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/009,832
FILING DATE: 01-Jan-1996
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Ph.D., Kathleen M.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 345-9100
TELEFAX: (617) 345-9111
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-726-306A-28
Alignment Scores:
Pred. No.: 1,11e-84 Length: 660
Score: 972.00 Matches: 196
Percent Similarity: 98.49% Conservative: 3
Best Local Similarity: 98.49% Mismatches: 0
Query Match: 61.48% Indels: 0
Gaps: 0
US-09-827-854-14 (1-317) x US-08-726-306A-28 (1-660)
Oy 1 MetLysValLeuThrProAlaLeuLeuValTrpPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAAGCTTCTGTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 121
Oy 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 122 GAGCAAGCGGTGGAGAGACAGAGCGGAGCGGAGCTGCGGCACAGAGAGAGAGAGAGAG 181
Oy 41 GlyGlnArgTrpGluLeuAlaLeuGluArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 182 GCCCAGCGCTGGAGACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 241
Oy 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValTrpGlnGluLeuArgAla 80
Db 242 CTGTCTGAGCAGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 301
Oy 81 LeuMetAspGluThrMetLysGluLeuLysValTyrLysSerGluLeuGluGluGluGlu 100
Db 302 CTGATGAGAGAGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 361
Oy 101 ThrProValAlaGluGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGln 120
Db 362 ACCCGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 421
Oy 121 ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 422 CGGCTGGCGCGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481

```


QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGlnLeuArgValArgLeuGlnAlaSerHisLeuArg 160
Db 482 CAGGCCATGCTCGGCGACAGACCGGAGAGACTGGGGTGCGCTCGCTCCACCTGGCGG 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AAGCTGGCTAAGCGCGCTCTCCGCGATCCCGATGACCTGCGAAGAGCCCTGGCGAGTGTAC 601
QY 181 GlnAlaGlyAlaArgGlnGlyAlaGluArgGlyLeuSerAlaIleArgGlnArgLeu 199
Db 602 CAGCGCGGGGCGCCGCGAGGGCGCCGAGCGGGGCTCAGCGCCATCCGCGAGCGCGCTG 658

RESULT 7
US-07-849-389-6
; Sequence 6, Application US/07849389
; Patent No. 5525493

APPLICANT: HORNES, Erik
 APPLICANT: UHLEN, Mathias
 TITLE OF INVENTION: CLONING METHOD AND KIT
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:

STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22313-0299

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 05/07/849,389

```

; FILING DATE: 19920519
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29, 768
REFERENCE/DOCKET NUMBER: 16787/168/DEBC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109

TELEX: 899149
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 base pairs

LENGTH: 330 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 ORGANISM: HUMAN TIDOCOCYTES
 CELL

```

ORGANISM: HUMAN LIPOPROTEIN E GENE
FEATURE:
NAME/KEY: CDS
LOCATION: 1 330

```

US-07-849-389-6

Alignment Scores:	
Pred. No.:	1.05e-41
Score:	519.00
Percent Similarity:	99.19%
Best Local Similarity:	98.08%
Query Match:	32.83%
DB:	1
Length:	330
Matches:	108
Conservative:	1
Mismatches:	1
Indels:	0
Gaps:	0

US-09-827-854-14 (1-317) x US-07-849-389-6 (1-330)

QY 87 LysIleuLeuLysAlaTyrIleuSerGluLeuGluIleuThrProValAlaGlu 106
|||||
1 AAGACGTGAAGGCGCTCAATCGGAACCTGACCCCGGTGGCGGAGAC 60

Oy	107	ThArGAlaAqLeuSerLySGluLeuGlnAlaIaGlnAlaArqLeuGlyAlaAspMet	126
Db	61	ACGGGGACACGGCTGTCTCCAGGACCTGCAGGGCGGAGGGCCCCGCTGGCGCGACATG	120
Oy	127	GIuAspValArGjIyArqLeuValGlnTyArqGIyGluValGlnAlaMetLeuGIyGln	146
Db	121	GAGGACGTGGCGCGCGCTGTCTGCAGTACGGCGGACAGGTGCAGGCATCTGGCCAG	180
Oy	147	SerPTrGluGluLeuArqValArqLeuAlaSerHisLeuArqLyLeuArqLyArqLeu	166
Db	181	AGCACCGAGGACTGCGGGTGGCGCTCGCTCCCACTGGCGGAAGCTGCGTAAAGGGCTC	240
Oy	167	LeuArqAspAlaAspAspLeuGlnLyArqLeuAlaValTyrglnAlaGlyAlaArqGlu	186
Db	241	CTCCGCGATGCCGATGACTCAGAAAGCGGCTCGGCACTGATACCGAGCGGGGCGCGGAG	300
Oy	187	GlyAlaGluArqGlyLeuSerAlaIleArq	196
Db	301	GGCGCCGAGCGCGGCTCTAGCGCCATCCG	330

RESULT 8
US-08-617-356-24
; Sequence 24, Application US/08617256
; Patent No. 6043031
GENERAL INFORMATION:

GENERAL INFORMATION:
APPLICANT: Kuster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
ATTORNEY: J. H. COCKFIELD, JR., 1000 RIVERSIDE DRIVE, SUITE 100, RIVERSIDE, CA 92506, U.S.A.

SIRENED: 00 State Street, Suite 510
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02109-1875
 COMPUTER READABLE FORM:

```

? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/617,256
?

```

; FILING DATE: March 18, 1996
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; ADDITIONAL NUMBER: 08/406 100

APPLICATION NUMBER: 08/400,193
FILING DATE: March 17, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth A.

REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: SQI-013CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid

```

```

; STRANDEDNESS: sin
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-617-256-24

```

Alignment Scores:	
Pred. No.:	6, 32e-27
Score:	363.00
Percent Similarity:	96.43%
Best local Similarity:	96.43%
Query Match:	22.96%
DB:	3
AB:	3
Length:	252
Matches:	81
Conservative:	0
Mismatches:	3
Indels:	0
Gaps:	0

US-09-827-854-14 (1-317) x US-08-617-256-24 (1-252)

QY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
DB 2 GCACGGCTGCTCCAGAGAGCTGCAGGGCGCCGACGCCGCTGGCGCCGACATGAGAGAC 61
QY 129 ValArgGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GGTGTG-CGCGCGCTGTGTGAGTACCGCGGAGGTGTCAGGCCATGCTCGCGCCAGACACCC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGGGGTGGCTGCCTCCACCTGCGCAAGCTGGTAAGCGGCTCTCCGG 180
QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAAAGTCCCTGGCACTGTACAGGCCGCGGCGCGAGAGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTCTC 252
RESULT 9
US-09-287-141-24
Sequence 24, Application US/09287141
Patent No. 6197498
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,141
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-287-141-24
Alignment Scores:
Pred. No.: 6,32e-27 Length: 252
Score: 363.00 Matches: 81

Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 22.96% Indels: 1
DB: 4 Gaps: 0
US-09-827-854-14 (1-317) x US-09-287-141-24 (1-252)
QY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
DB 2 GCACGGCTGCTCCAGAGAGCTGCAGGGCGCCGACGCCGCTGGCGCCGACATGAGAGAC 61
QY 129 ValArgGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GGTGTG-CGCGCGCTGTGTGAGTACCGCGGAGGTGTCAGGCCATGCTCGCGCCAGACACCC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGGGGTGGCTGCCTCCACCTGCGCAAGCTGGTAAGCGGCTCTCCGG 180
QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAAAGTCCCTGGCACTGTACAGGCCGCGGCGCGAGAGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTCTC 252
RESULT 10
US-09-431-613-24
Sequence 24, Application US/09431613
Patent No. 6221601
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/431,613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002G
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: CDNA
US-09-431-613-24

Alignment Scores:
Pred. No.: 6,32e-27 Length: 252
Score: 363.00 Matches: 81
Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 22.96% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-14 (1-317) x US-09-431-613-24 (1-252)

OY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
DB 2 GCACGGCTGTCCAGAGACTGCAGCGCGCGCGCGCGCGCGCGCATGAGAGAC 61
OY 129 ValArgGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTG-CGCCCGCTGTGCAGTACCGCGCGCGCGCGCGCGCATGCTCGGCAGAGCAC 120
OY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGCGGGTGCCTGCGCTCCACCTGCGCAAGCTGCTAACCGGCTCTCCGC 180
OY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAACTCCCTGCACTGTACAGCGCGCGCGCGCGCGCGCC 240
OY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCGCTC 252

RESULT 11

US-09-504-245-24
Sequence 24, Application US/09504245
Patent No. 6221605

GENERAL INFORMATION:

APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: Heller Ehrman White & McCaulliffe LLP
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037-9103

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/504,245

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002J
TELECOMMUNICATION INFORMATION:
TELEPHONE: 858-450-8400
TELEFAX: 858-587-5360
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-504-245-24

Alignment Scores:
Pred. No.: 6,32e-27 Length: 252
Score: 363.00 Matches: 81
Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 22.96% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-14 (1-317) x US-09-504-245-24 (1-252)

OY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
DB 2 GCACGGCTGTCCAGAGACTGCAGCGCGCGCGCGCGCGCGCATGAGAGAC 61
OY 129 ValArgGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTG-CGCCCGCTGTGCAGTACCGCGCGCGCGCGCGCGCATGCTCGGCAGAGCAC 120
OY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGCGGGTGCCTGCGCTCCACCTGCGCAAGCTGCTAACCGGCTCTCCGC 180
OY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAACTCCCTGCACTGTACAGCGCGCGCGCGCGCGCGCC 240
OY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCGCTC 252

RESULT 12

US-09-287-682-24
Sequence 24, Application US/09287682
Patent No. 6235478

GENERAL INFORMATION:

APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.

TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: Heller Ehrman White & McCaulliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037-9103

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/287,682

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002E
TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-287-682-24

Alignment Scores:
Pred. No.: 6,32e-27 Length: 252
Score: 363.00 Matches: 81
Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 22.96% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-14 (1-317) x US-09-287-682-24 (1-252)

QY 109 AlaArgLeuSerIysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
|||||
DB 2 GCACGGCTCTCCAGAGAGCTGCAGCGCGCGAGCCGCGCTGCGCGACATGAGAGAC 61

QY 129 ValArgGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
|||||
DB 62 GTGTG-CGCGGCTGTGTGAGTACCGCGGCGAGGTGCAGGCCATGCTCGGCCAGACACC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
|||||
DB 121 GAGGAGCTCGGGTGGCTCGCTCCGCTCCACCTCCGCAAGCTGCTAAGCGGCTCTCCGC 180

QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
|||||
DB 181 GATGCCGATGACTGCAGAAATCCCTGGCAGTGTACAGGCGGGGCGCGCGAGGGCGCC 240

QY 189 GluArgGlyLeu 192
|||||
DB 241 GAGCGGGGCTC 252

RESULT 13
US-09-287-679-24
Sequence 24, Application US/09287679
Patent No. 6258538
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
APPLICANT: Braun, Andreas
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McAuliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,679
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-287-679-24

Alignment Scores:
Pred. No.: 6,32e-27 Length: 252
Score: 363.00 Matches: 81
Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 22.96% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-14 (1-317) x US-09-287-679-24 (1-252)

QY 109 AlaArgLeuSerIysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
|||||
DB 2 GCACGGCTCTCCAGAGAGCTGCAGCGCGCGAGCCGCGCTGCGCGACATGAGAGAC 61

QY 129 ValArgGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
|||||
DB 62 GTGTG-CGCGGCTGTGTGAGTACCGCGGCGAGGTGCAGGCCATGCTCGGCCAGACACC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
|||||
DB 121 GAGGAGCTCGGGTGGCTCGCTCCGCTCCACCTCCGCAAGCTGCTAAGCGGCTCTCCGC 180

QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
|||||
DB 181 GATGCCGATGACTGCAGAAATCCCTGGCAGTGTACAGGCGGGGCGCGCGAGGGCGCC 240

QY 189 GluArgGlyLeu 192
|||||
DB 241 GAGCGGGGCTC 252

RESULT 14
US-09-397-766-24
Sequence 24, Application US/09397766
Patent No. 6268144
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
APPLICANT: Braun, Andreas
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McAuliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:


```

1 APPLICATION NUMBER: US/09/397,766
2 FILING DATE:
3 CLASSIFICATION:
4 PRIOR APPLICATION DATA:
5 APPLICATION NUMBER: 08/617,256
6 FILING DATE: 18-MAR-1996
7 PRIOR APPLICATION DATA:
8 APPLICATION NUMBER: 08/406,199
9 FILING DATE: 03-MAR-1995
10 ATTORNEY/AGENT INFORMATION:
11 NAME: Seidman, Stephanie L
12 REGISTRATION NUMBER: 33,779
13 REFERENCE/DOCKET NUMBER: 24736-2002H
14 TELECOMMUNICATION INFORMATION:
15 TELEPHONE: 619-450-8400
16 TELEFAX: 619-587-5360
17 TELECOMMUNICATION INFORMATION:
18 TELEPHONE: (617)227-7400
19 TELEFAX: (617)227-5941
20 INFORMATION FOR SEQ ID NO: 24:
21 SEQUENCE CHARACTERISTICS:
22 LENGTH: 252 base pairs
23 TYPE: nucleic acid
24 STRANDEDNESS: single
25 TOPOLOGY: linear
26 MOLECULE TYPE: CDNA
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
98
```


GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 17:32:58 ; Search time 65.8753 Seconds
(without alignments)
3377.460 Million cell updates/sec

Title: US-09-827-854-14

Perfect score: 1581

Sequence: 1 MKVLMALVTLFLAGCOAKV.....VEKVOAAGTSAPVPSDNH 317

Scoring table:

BIOSUM62	
Xgapop 10.0 , Ygapext 0.5	
Xgapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 501302 seqs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp
-O=/cgn2_1/USPfo.spool/US09827854/r/unat_11032003_101612_27578/app_query.fasta.1.3576
-DB=published.Applications_NA -OFMT=fastap -SUFFIX=p2n.rmpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blcosum62
-TRANS=human40.cdi -LIST=45 -DOCCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09827854.ecgn_1.1.298.e/unat_11032003_101612_27578
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_WMAP -LARGEXQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=130 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -RGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCIT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/PCITUS_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1581	100.0	1156	10	US-09-827-854-7
2	1573	99.5	1156	9	US-09-870-759-129
3	1573	99.5	1156	9	US-09-802-640-17
4	1573	99.5	1156	10	US-09-827-854-8

5	1573	99.5	1291	12	US-10-044-090-454	Sequence 454, App
6	1569	99.2	1156	10	US-09-827-854-12	Sequence 12, Appl
7	1565	99.0	1156	10	US-09-827-854-9	Sequence 9, Appl
8	1565	99.0	1156	10	US-09-827-854-11	Sequence 11, Appl
9	1564	98.9	1157	10	US-09-934-456-760	Sequence 760, App
10	1564	98.9	1157	10	US-09-880-107-2244	Sequence 2244, App
11	1558	98.5	1156	10	US-09-827-854-10	Sequence 10, Appl
12	1370.5	86.7	41907	10	US-09-967-013-5	Sequence 5, Appl
13	867.5	54.9	786	10	US-09-925-902-133	Sequence 133, App
14	625	39.5	478	10	US-09-964-824-374	Sequence 374, App
15	625	39.5	478	10	US-09-880-107-2491	Sequence 2491, App
16	471	29.8	356	10	US-09-960-352-5420	Sequence 5420, App
17	466.5	29.5	449	10	US-09-960-352-4726	Sequence 4726, App
18	443	28.0	414	10	US-09-960-352-4237	Sequence 4237, App
19	439.5	27.8	442	10	US-09-960-352-9395	Sequence 9395, App
20	428.5	27.1	416	10	US-09-960-352-8720	Sequence 8720, App
21	422.5	26.7	423	10	US-09-960-352-14047	Sequence 14047, A
22	405.5	25.6	409	10	US-09-960-352-5148	Sequence 5148, App
23	401.5	25.4	416	10	US-09-960-352-9797	Sequence 9797, App
24	389	24.6	253	10	US-09-179-5368-130	Sequence 130, App
25	387.5	24.5	425	10	US-09-960-352-3497	Sequence 3497, App
26	380.5	24.1	390	10	US-09-960-352-1311	Sequence 1311, App
27	379.5	24.0	391	10	US-09-960-352-1278	Sequence 1278, App
28	379.5	24.0	392	10	US-09-960-352-10599	Sequence 10599, A
29	379.5	24.0	401	10	US-09-960-352-4187	Sequence 4187, App
30	375.5	23.8	413	10	US-09-960-352-8042	Sequence 8042, App
31	369.5	23.4	404	10	US-09-960-352-2325	Sequence 2325, App
32	369.5	23.4	405	10	US-09-960-352-14063	Sequence 14063, A
33	363	23.0	252	10	US-09-796-416-24	Sequence 24, Appl
34	363	23.0	252	10	US-09-879-341-24	Sequence 24, Appl
35	351.5	22.2	377	10	US-09-960-352-176	Sequence 176, App
36	350.5	22.2	353	10	US-09-960-352-4914	Sequence 4914, App
37	350.5	22.2	378	10	US-09-960-352-10040	Sequence 10040, A
38	350.5	22.2	386	10	US-09-960-352-10361	Sequence 10361, A
39	348.5	22.0	388	10	US-09-960-352-11986	Sequence 11986, A
40	344.5	21.9	396	10	US-09-960-352-1186	Sequence 3457, App
41	344.5	21.8	377	10	US-09-960-352-10540	Sequence 10540, A
42	338.5	21.4	432	10	US-09-960-352-8934	Sequence 8934, App
43	338	21.4	332	10	US-09-960-352-6552	Sequence 6552, App
44	337.5	21.3	376	10	US-09-960-352-5408	Sequence 5408, App
45	335.5	21.2	373	10	US-09-960-352-3681	Sequence 3681, App

ALIGNMENTS

RESULT 1
US-09-827-854-7
Sequence 7, Application US/09827854
Patent No. US20020123093A1
GENERAL INFORMATION:
APPLICANT: Zannis, Vassilis
APPLICANT: Kyriacs, Kyriakos E.
TITLE OF INVENTION: Compounds and methods for lowering
FILE REFERENCE: 07180/004003
CURRENT APPLICATION NUMBER: US/09/827,854
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: US 09/679,088
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 09/544,386
PRIOR FILING DATE: 2000-04-06
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 1156
TYPE: DNA
ORGANISM: Homo sapiens
US-09-827-854-7
Alignment Scores:
Pred. No.: 7.4e-134
Score: 1581.00
Percent Similarity: 100.00%
Length: 1156
Matches: 317
Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-09-827-854-14 (1-317) x US-09-827-854-7 (1-1156)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||||||
 Db 61 ATGAAGTTCTGTGGCTGCTGTGTCACATTCTCTGCAGAGATGCCAGGCAAGGTG 120

QY 21 GluGlnAlaValaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 |||||||
 Db 121 GAGCAAGCGGTGAGAGAGCCGAGCCGAGCTGGCCAGAGACCGAGTGCAGAGC 180

QY 41 GlyLlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 |||||||
 Db 181 GGCAGAGCTGTGGAGACTGGAGCTGGCTTTTGGGATTACTGGGCTGGTGGAGACA 240

QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||||||
 Db 241 CTGTCTGACGAGGTGACGAGGAGCTGCTCAGCTCCAGGTACCCAGGAACTGAGGGCG 300

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 |||||||
 Db 301 CTGATGACGAGACCATGAAGAGTTGAAGCCTTACAAATCGGAACTGGAGAACTG 360

QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
 |||||||
 Db 361 ACCCGGTGGCGAGAGAGAGCGGCGACGCTCTCCAGAGAGCTGCAGCGGGCGAGGCC 420

QY 121 ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyGluVal 140
 |||||||
 Db 421 CGGCTGGGCGGACATGAGAGAGCTGGCGGCGGCTGTGCATGACCGCGGAGAGTG 480

QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValaArgLeuAlaSerHisLeuArg 160
 |||||||
 Db 481 CAGGCCATGCTCGCCAGAGACCGAGAGCTCGGGTGGCTCCGCTCCACCTGGCG 540

QY 161 LysLeuAArgLysArgLeuAArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||||||
 Db 541 AAGCTGGTAAAGGGGCTCTCCGCGATGCCGATGACCTGAGAAAGGGCTGGAGATGAC 600

QY 181 GlnAlaGlyAlaArgGluGluGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 |||||||
 Db 601 CAGGCGGGGCGCGAGGGCGCGAGCGGCTCAGCGCCATCCGCGAGCGCTGGGG 660

QY 201 ProLeuValGluGlnGlyArgValaArgAlaAlaThrValGlySerLeuAlaGlnPro 220
 |||||||
 Db 661 CCCCTGGTGAACAGGGCCGCTGGCGGCGCCACTGTGGGCTCCCTGGCGGCGAGCG 720

QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
 |||||||
 Db 721 CTACAGAGAGGCGCCAGGCTGGGGGAGCGGCTGGCGCGCGGATGGAGAGATGGCG 780

QY 241 SerTrpThrArgAspArgLeuAspGluValLysGluGlnValaGluValaArgAlaLys 260
 |||||||
 Db 781 AGCGGACCGCGACCGCTCGAGCGAGGTGAAGAGACAGAGTGGCGAGGTGGCGCAG 840

QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 |||||||
 Db 841 CTGGAGAGAGAGCCAGCGATATCGCTTCAGCGGAGCGCTTCCAGAGCGCCCTCAG 900

QY 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 |||||||
 Db 901 AAGTGTGTGAGAGCCCTGTGGTGAAGACATGACGCGCAGTGGGCGCGGCTGGTGGAGAG 960

QY 301 ValGlnAlaAlaValaGlyThrSerAlaAlaProValProSerAspAsnHis 317
 |||||||
 Db 961 GTGACAGCTGCGTGGGACACAGAGCGCGCCCTGTGCTGCCAGCAGCAATATCAC 1011

RESULT 2
 US-09-870-759-129

; Sequence 129, Application US/09870759

; Patent No. US20020177551A1
 ; GENERAL INFORMATION:
 ; APPLICANT: TERMAN, David S
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
 ; FILE REFERENCE: 870759
 ; CURRENT APPLICATION NUMBER: US/09/870,759
 ; PRIORITY FILING DATE: 2002-01-14
 ; PRIOR APPLICATION NUMBER: US 60/208,128
 ; NUMBER OF SEQ ID NOS: 166
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 129
 ; LENGTH: 1156
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (61)..(1014)
 ; OTHER INFORMATION:
 ; US-09-870-759-129

Alignment Scores:
 Pred. No.: 3,88e-133 Length: 1156
 Score: 1573.00 Matches: 316
 Percent Similarity: 99.68% Conservative: 0
 Best Local Similarity: 99.68% Mismatches: 1
 Query Match: 99.49% Indels: 0
 DB: 9 Gaps: 0

US-09-827-854-14 (1-317) x US-09-870-759-129 (1-1156)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||||||
 Db 61 ATGAAGTTCTGTGGCTGCTGTGTCACATTCTCTGCAGAGATGCCAGGCAAGGTG 120

QY 21 GluGlnAlaValaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 |||||||
 Db 121 GAGCAAGCGGTGAGAGAGCCGAGCCGAGCTGGCCAGAGACCGAGTGCAGAGC 180

QY 41 GlyLlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 |||||||
 Db 181 GGCAGAGCTGTGGAGACTGGAGCTGGCTTTTGGGATTACTGGGCTGGTGGAGACA 240

QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||||||
 Db 241 CTGTCTGACGAGGTGACGAGGAGCTGCTCAGCTCCAGGTACCCAGGAACTGAGGGCG 300

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 |||||||
 Db 301 CTGATGACGAGACCATGAAGAGTTGAAGCCTTACAAATCGGAACTGGAGAACTG 360

QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
 |||||||
 Db 361 ACCCGGTGGCGAGAGAGAGCGGCGACGCTGTCCAAAGAGCTGCAGCGGGCGAGGCC 420

QY 121 ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyGluVal 140
 |||||||
 Db 421 CGGCTGGGCGGACATGAGAGAGCTGGCGGCGGCTGTGCATGACCGCGGAGAGTG 480

QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValaArgLeuAlaSerHisLeuArg 160
 |||||||
 Db 481 CAGGCCATGCTCGCCAGAGACCGAGAGCTGGCGGCTGGCTCCAGGCTCCAGCTGGCG 540

QY 161 LysLeuAArgLysArgLeuAArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||||||
 Db 541 AAGCTGGTAAAGGGGCTCTCCGCGATGCCGATGACCTGAGAAAGGGCTGGAGATGAC 600

QY 181 GlnAlaGlyAlaArgGluGluGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 |||||||
 Db 601 CAGGCGGGGCGCGAGGGCGCGAGCGGCTCAGCGCCATCCGCGAGCGCTGGGG 660

QY 201 ProLeuValGluGlnGlyArgValaArgAlaAlaThrValGlySerLeuAlaGlnPro 220
 |||||||

Db 661 CCCCTGGTGAACAGAGGCCCGCTGGCGGCCGCTGAGCTCCCTGGCCGCCAGCCG 720
QY 221 LeuGlnGluArgAlaGlnAlaTrrpGlyGluArgLeuArgAlaArgMetGlnGlnMetGly 240
Db 721 CTACAGAGAGCGGCGCCAGGCTGGGGCGGAGCGGCTGGCGCGCGGATGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValAlaArgAlaLys 260
Db 781 ACCCGAGACCGCGGACCGCTGGACGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 841 CTGGAG 900
QY 281 SerTrpPheGluProLeuValGlnLysPheGlnArgGlnTrpAlaGlnGlyLeuValGlnLys 300
Db 901 ACCTGTTTCAGACCCCTGGTGGAAACATGACAGCCAGCTGGGGCGCGGCTGGTGGAGAG 960
QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCCGTGGGACACAGCCCGCCCTGTGCTCCAGAGACATTCAC 1011
RESULT 3
US-09-802-640-17
; Sequence 17, Application US/09802640
; Publication No. US20030036057A1
; GENERAL INFORMATION:
; APPLICANT: Braun, Andreas
; APPLICANT: Bonsai Aruna
; APPLICANT: Kleya Patrick
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND THEIR USE
; FILE REFERENCE: 24736-2048
; CURRENT APPLICATION NUMBER: US/09/802, 640
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)...(1014)
; OTHER INFORMATION: Nucleotide Sequence encoding apolipoprotein E
; OTHER INFORMATION: (APOE)
US-09-802-640-17
Alignment Scores:
Pred. No.: 3,88e-133 Length: 1156
Score: 1573.00 Matches: 316
Percent Similarity: 99.68% Conservatave: 0
Best Local Similarity: 99.68% Mismatches: 1
Query Match: 99.49% Indels: 0
DB: 9 Gaps: 0
US-09-827-854-14 (1-317) x US-09-802-640-17 (1-1156)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGGTCTCTGTGGCTGCGCTTGTCTGTCATCTCGGAGATGCCAGGCCAAGGTG 120
QY 21 GlnGlnAlaValAlaGlnThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAAGGGGTGAGACAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 180
QY 41 GlnGlnArgTrpGlnLeuAlaLeuGlnLysArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 181 GGCCAGCGCTGGAGACTGGCTGCTTTGGGATTTACTCTGGCTGGGTGAGAGCA 240
QY 61 LeuSerGlnGlnValAlaGlnGlnLysLeuLeuSerSerGlnValThrGlnGlnLysValAla 80
Db 241 CTGTCTGAGCAGAGTGCAGAGAGAGAGTGTCTCCAGGTCCAGGACCCAGAGACTGAGGCG 300

QY 81 LeuMetAspGluThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
Db 301 CTGATGAGAGAGACCATGAGAGAGTTGAAGGCTTCAATTCGAGACTGGAGAACACTG 360
QY 101 ThrProValAlaGlnGlnLysThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
Db 361 ACCCGGTGGGGAGAGAGAGCGGCGGAGCGGCTGTCTCAAGAGAGCTCAGAGCGCGCAGCC 420
QY 121 ArgLeuGlnAlaAspMetGlnAspValArgGlyArgLeuValGlnTrpArgGlyGlnVal 140
Db 421 CGGCTGGGCGGAGACATGAGAGAGCTGTGGCGCCCTGTGTCTACATACCGGCGAGGTG 480
QY 141 GlnAlaMetLeuGlnLysSerThrGlnLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCTATGCTGGCCAG 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTGGCTGAAGCGGCTCTCTCGGATGCCATGACCTGCAAGAGCGCTGGCACTGTAC 600
QY 181 GlnAlaGlnAlaArgGlnGlnAlaGlnLysLeuSerAlaIleArgGlnArgLeuGly 200
Db 601 CAGCGCGGGGCGCGGAGAGCGCGGAGCGCGCTCAGCGCATCCGAGAGCGCTGGGG 660
QY 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlnSerLeuAlaGlyGlnPro 220
Db 661 CCCCTGGTGAACAGAGGCCGCTGGCGCGCCGCTGTGTGGCTCTCTGGCGCGGAGCGG 720
QY 221 LeuGlnGluArgAlaGlnAlaTrrpGlyGluArgLeuArgAlaArgMetGlnGlnMetGly 240
Db 721 CTACAGAGAGCGGCGCCAGGCTGGGGCGAGCGGCTGGCGCGCGGATGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValAlaArgAlaLys 260
Db 781 AGCGGAGACCGCGGACCGCTGGAGAGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 841 CTGGAG 900
QY 281 SerTrpPheGluProLeuValGlnLysPheGlnArgGlnTrpAlaGlnGlyLeuValGlnLys 300
Db 901 AGCTGTTTCAGACCCCTGGTGGAAACATGACAGCCAGCTGGGGCGGCTGGTGGAGAGAG 960
QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCCGTGGGACACAGCCCGCCCTGTGCTCCAGAGACATTCAC 1011
RESULT 4
US-09-827-854-8
; Sequence 8, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Kyriacos, Kyriacos F.
; APPLICANT: Kyriacos, Kyriacos F.
; TITLE OF INVENTION: Compounds and methods for lowering
; TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827, 854
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679, 088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544, 386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-8

Db 726 CAGCCGCGGCGCCGAGGCGCGCCGAGCGCCCTCAGCGCCATCCGAGCGCCCTGGG 785
Qy 201 ProleuValuGlugIngluValArgValAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 786 CCCCTGGTGGAAAGAGGCGCGCTGGGGCGCCACTGTGGCTCCCTGGCCGCGCAAGCG 845
Qy 221 LeugIngluArgAlaGlnAlaThrPgluArgLeuArgAlaArgMetGluGlnMetGly 240
Db 846 CTACAGAGAGCGGCGCCAGGCGCTGGGGCGAGCGCGCTGGCGCGCGATGAGAGATGGCG 905
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 906 ACCCGAGCGCGCGAGCGCGCTGGAGCGAGTGGCGAGGTGGCGAGGTGGCGCGCAAG 965
Qy 261 LeugIngluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 966 CTGGAGAGAGCGCGCCAGCATGAGCTGGAGCGCGAGGCTTCAGAGCGCGCTCGCAAG 1025
Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 1026 ACCTGGTTCGAGCGCCCTGGTGGAAAGACATGCAGCGCCAGTGGCGCGCTGGTGGAGAAG 1085
Qy 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 1086 GTGCAGGCTGCGCGTGGGACACGAGCGCGCCCTGTGCCAGCGACATAC 1136

RESULT 6

US-09-827-854-12
; Sequence 12, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassiliis
; APPLICANT: Kyriacos, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827, 854
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679, 088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544, 386
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-12

Alignment Scores:

Pred. No.: 8,899-133 Length: 1156
Score: 1569.00 Matches: 315
Percent Similarity: 99.68% Conservative: 1
Best Local Similarity: 99.37% Mismatches: 1
Query Match: 99.24% Indels: 0
DB: 10 Gaps: 0

US-09-827-854-14 (1-317) x US-09-827-854-12 (1-1156)

Qy 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGTTCTGTGGCTGGCGTCTGGTGCATCTCGGCGAGATGCGAGCCCAAGTGG 120
Qy 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAAGCGGTGGAGACAGAGCGCGAGCGCGCGCCAGACAGACGAGTGGCGAGCG 180
Qy 41 GlyGlnArgTrpGluLeuAlaLeuGlnValArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 181 GGGCAGCGCGTGGAGACATGCGACTGGGTGGCTTTTGGGATTCACCTGCGTGGCGAGCA 240
Qy 61 LeuSerGlnGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80

Db 241 CTGCTGAGACAGGTGCAGAGAGAGCTGCTCAGCTCCAGGTCCACAGAACTGAGGGCG 300
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
Db 301 CTGATGAGACGAGACCATGATGAGAGTGAAGGCTCCAAATGGAAGTGGAGAACAACTG 360
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 361 ACCCGGTGGCGAGAGAGAGCGGGGACGAGCTGTCCAAAGAGACTTCAGGGCGGCGAGGCC 420
Qy 121 ArgLeuGlnAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyLysVal 140
Db 421 CGCGTGGCGCGGACATGAGAGAGCTGTGCGCGCGCTGTGGTCACTACCGCGGCGAGGTG 480
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCGATGCTCGGCGACAGACACCGAGAGAGCTGCGGGTGGCGCTCCCTCCACCTGGCG 540
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGTGGCTCAGCGGCTCTCCGCGATGCCATGACCTGCGAGAACGCGCTGGAGTGTAC 600
Qy 181 GlnAlaGlyAlaArgGluGlyValAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 601 CAGCGCGGGCGCGCGAGGCGCGCGAGCGCGGCTCAGCGCATCCGCGAGCGCGCTGGGG 660
Qy 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGGTGGAAAGAGGCGCGGTGGGGCGCGCTGTGGGTCTCCCTGGCGGCGAGCGCG 720
Qy 221 LeugIngluArgAlaGlnAlaThrPgluArgLeuValArgAlaArgMetGluGlnMetGly 240
Db 721 CTACAGAGAGCGGCGCCAGGCTGGGGCGAGCGCGCTGGCGCGGATGAGAGATGGCGC 780
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 781 AGCGGAGCGCGGAGCGCGCTGGAGCAGGTGAAGAGAGAGGTGGCGGAGTGGCGCGCAAG 840
Qy 261 LeugIngluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 841 CTGGAGAGAGCGCGCCAGCATGAGTGCAGCGCGCGAGGCTTCAGAGCGCGCTCGCAAG 900
Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 901 ACCTGGTTCGAGCGCCCTGGTGGAAAGACATGCGAGCGCGAGTGGCGCGCTGGTGGAGAAG 960
Qy 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCGGTGGGACACGAGCGCGCCCTGTGCCAGCGACATAC 1011

RESULT 7

US-09-827-854-9
; Sequence 9, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassiliis
; APPLICANT: Kyriacos, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827, 854
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679, 088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544, 386
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens


```

; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherif, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2244
; LENGTH: 1157
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession NO. US20020142981A1 M12529
US-09-880-107-2244

```

```

Alignment Scores:
Pred. No.: 2,51e-132 Length: 1157
Score: 1564.00 Matches: 314
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 99.05% Mismatches: 3
Query Match: 98.92% Indels: 0
DB: Gaps: 0

```

US-09-827-854-14 (1-317) x US-09-880-107-2244 (1-1157)

```

QY 1 MetLysValLeuTrpAlaIleuValTrpPheLeuAlaGlyCysGlnAlaLysVal 20
DB 62 ATGAAGGTCTCTGGGCTGGCTGGTCTGTCACATTCTGCGACAGATGCCAGGCAAGTG 121
QY 21 GlnGlnAlaValaGlnThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 122 GACCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCGCAGACAGCCGAGTGAGAGC 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 182 GGCACAGCGCGGAGACTGGCTGGTGGCTGGTGGATTAAGCTGGCTGGCTGGAGACA 241
QY 61 LeuSerGlnGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 242 CTGTCTGAGAGAGGTGACAGAGAGAGCTGCTCAGCTCCCAAGTCCCAAGACGAGAGCGC 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGlnLeuGlnGlnGlnLeu 100
DB 302 CTGATGAGAGAGACCAATGAAGAGTTGAAGGCTTACCAAAATCGGAACCTGGAGCAACTG 361
QY 101 ThrProValAlaGlnGluThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
DB 362 ACCCGGTGAGGAGAGAGAGCGGCGACGCTGTCCAAGAGCTGCAGACGGCGCAGGCC 421
QY 121 ArgLeuGlnAlaAspMetGluAspValArgGlyArgLeuValGlnThrArgGlyGlnVal 140
DB 422 CGGCTGGCGCGGACATGAGAGAGCTGTGCGCGCGCTGTGTGACAGTACCAGCGGAGAGTG 481
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 482 CAGGCGCATCTGGCGCAGAGACCGAGAGCTGCGGGTGGCGCTGCCCTGCCACTGCGCG 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 542 AAGCTGCGTGAAGGCGCTCTCCGCGCATCCGATCCGAGACCTCGAGAAAGCGCTGGCAGTGC 601
QY 181 GlnAlaGlyAlaArgGlnGlnAlaArgGlyLysSerAlaIleArgGlnArgGlnGly 200
DB 602 CAGCGCGGCGCGCGAGAGCGCGGAGCGCGGCTCAGCGCCCTCCGCGAGCGCGCTGGGG 661
QY 201 ProLeuValGlnGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220

```

```

DB 662 CCCCTGGTGAGACAGGCGCGCTGCGGCGCCACACTGTGGCTCCCTGCGCGGCGAGCGG 721
QY 221 LeuGlnGlnArgAlaGlnAlaTrpGlyArgLeuArgAlaArgMetGlnGlnMetGly 240
DB 722 CTACAGAGAGCGGCGCCAGCGCTGGGCGAGCGGCTCGCGCGGAGTGAAGAGATGGGC 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValaGlnAlaValArgAlaLys 260
DB 782 AGTCGAGCCCGCGAGCGCTCGAGCGAGGTGAAGAGACAGTGGCGGAGGTGCGCGCGCAAG 841
QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 842 CTGAGAGAGAGCGCCACAGAGATACCTGACAGCGCGAGCGCTTCCAGGCGCGCTCAAG 901
QY 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlnLeuValGlnLys 300
DB 902 AGCTGGTTCAGGCGCTCGGTGGAGACATGACAGCGCAGTGGCGCGGCTGTGGAGAG 961
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 962 GTGCAGGCTGCGGTGGGCAACAGCGCGCGCTGTGCCAGGCAATCAC 1012

```

```

RESULT 11
US-09-827-854-10
; Sequence 10, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassilis
; APPLICANT: Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-10

```

```

Alignment Scores:
Pred. No.: 8,68e-132 Length: 1156
Score: 1558.00 Matches: 314
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 99.05% Mismatches: 3
Query Match: 98.55% Indels: 0
DB: Gaps: 0

```

US-09-827-854-14 (1-317) x US-09-827-854-10 (1-1156)

```

QY 1 MetLysValLeuTrpAlaIleuValTrpPheLeuAlaGlyCysGlnAlaLysVal 20
DB 61 ATGAAGGTCTCTGGGCTGGCTGGTCTGTCACATTCTGCGACAGATGCCAGGCAAGTG 120
QY 21 GlnGlnAlaValaGlnThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 121 GACCAAGCGGTGGAGACAGAGCCGAGAGCTGCGCAGACAGACGAGTGGCAGAGC 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 181 GGCACAGCGCTGGAACTGCACTGGCTGCTTTGGGATTAAGCTGCGCTGGCTGAGACA 240
QY 61 LeuSerGlnGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 241 CTGTCTGAGAGAGGTGACAGAGAGAGCTGTACAGTCCCAAGTCCCAAGAACTGAGAGCGC 300

```



```

QY 81 LeuMetaspGluThrMetLysGluLeuValAryLysSerGluLeuGluGluLeu 100
    |||||
Db 301 CTGATGAGACGACCATGAGAGAGTTGAAGGCTTCAATGCAATGCAACTGAGAGAACACTG 360
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGluAlaAlaGlnAla 120
    |||||
Db 361 ACCCGGGGGGAGAGACCGGGGCGCTGTCACAGAGAGCTCAGCGCGCGAGGCC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyValAla 140
    |||||
Db 421 CGGCGGGGGGAGACATGAGACCTGTGGCGCGCTGTGTGCATACCTCCGCGAGGTG 480
QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
    |||||
Db 481 CAGGCGCATGCTCGACCAAGACACCGAGAGAGCTGGCGGTGCCCTCGCTCCACTCGGC 540
QY 161 LysLeuArgGlyArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
    |||||
Db 541 AAGCTGCTAAGCGCTCTCCGCGATGCCATGCACCTCGAGAAAGTGGCTGGCAGTGTAC 600
QY 181 GlnAlaGlyAlaArgGluGluGlyAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
    |||||
Db 601 CAGGCGGGGGCGCGAGGGCGCGAGCGCGCTCAGCGCGCATCCGCGAGCGCTGGGG 660
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
    |||||
Db 661 CCCGTGTGTGAACAGAGGGCGGTGGGGCGCCCACTGTGGCTCTCGCGCGCGAGCGG 720
QY 221 LeuGlnGluArgAlaGlnAlaTyrGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
    |||||
Db 721 CTACAGAGCGGGCGCCAGCGCTGGGGCGAGCGCTGCGCGCGGATGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
    |||||
Db 781 AGCGGACCGCGGACCGCTGTGAGAGGTGAAGGAGCGGTGGCGGAGTGGCGCGCCAG 840
QY 261 LeuGlnGluGlnAlaGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
    |||||
Db 841 CTGGAGGAGCGAGCCCGACAGATACGCTGCAAGGCCGAGGCCCTTCAGGCCCGCTCAG 900
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTyrPalaglyLeuValGluLys 300
    |||||
Db 901 AGCTGTTCGAGCCCTGTGTGAAGACATGCAGCGCCAGTGGCGGCGCTGTGTGAGAG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
    |||||
Db 961 GTGAGGCTGCGGTGGGACACGCGCCCTGTGTGCCACGACAAATCAC 1011

RESULT 12
US-09-967-013-5
: Sequence 5, Application US/09967013
: Patent No. US2002005840A1
: GENERAL INFORMATION:
: APPLICANT: Stanton, Jr, Vincent P.
: TITLE OF INVENTION: METHOD FOR GENETIC ANALYSIS OF APOE DNA
: FILE REFERENCE: 11926-022001
: CURRENT APPLICATION NUMBER: US/09/967,013
: CURRENT FILING DATE: 2000-10-25
: PRIOR APPLICATION NUMBER: 60/206,613
: NUMBER OF SEQ ID NOS: 91
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 41907
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-967-013-5

Alignment Scores:
Pred. No.: 4,61e-113
Score: 1370.50
Percent Similarity: 60.00%
Length: 41907
Matches: 300
Conservative: 0

```

```

Best Local Similarity: 60.00%
Query Match: 86.69%
Mismatches: 6
Indels: 194
Gaps: 1
DB: 10

US-09-827-854-14 (1-317) x US-09-967-013-5 (1-41907)
QY 12 PheLeuAlaGlyCysGlnAlaValGluGlnAlaValGluThrGluProGluProGlu 31
    |||||
Db 20316 TTCCACACAGAGATGCCACRCCAGAGTGGAGACAGCGGTGGAGACAGAGCCGAGCCGAG 20375
QY 32 LeuArgGlnGlnThrGluTyrGlnSerGlyGlnArgTyrGluLeuAlaLeuGlyArgPhe 51
    |||||
Db 20376 CTGCGCCAGCAGACGAGTGGAGAGCGCCAGCGCTGGAACTGGCAGTGGGTGCTTTT 20435
QY 52 TrpAspTyrLeuArgTyrValGlnThrLeuSerGluGlnValGlnGluLeuSer 71
    |||||
Db 20436 TGGGATTACCTGGCGTGGGTGAGACACTGTCTGAGCAGGTGGAGAGAGACTGCTCAGC 20495
QY 72 SerGlnValThrGlnGluLeu-----
    |||||
Db 20496 TCCCAAGTCCACCCAGAGAACTGAGTGAAGTCCCATCTGCGCTTGACCTCTCGGTG 20555
QY 78 -----
Db 20556 GCGGCGTATACCTCCCGCAGGTTCATTTCAATCTGCCCCCTGTGCTAAGTCTTGGGGGG 20615
QY 78 -----
Db 20616 CCGTGGCTCTGCTGCTGTCTAGCTTCTTCCATTTCTGACTCTGAGTCTTGAAGTCTGC 20675
QY 78 -----
Db 20676 TGGAAATCTCTCTCTCAGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 20735
QY 78 -----
Db 20736 CGTCTGCGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 20795
QY 78 -----
Db 20796 CTCACGTGTGTGCCAGCGCTGTGTCTGAACTTCTGGGCTCAAGGATCTCCGCGCTCGG 20855
QY 78 -----
Db 20856 CCTCCCAAAGTGTGGGATTAGAGCATGAGCACCTTGGCGGCTCTTACTCTTCT 20915
QY 78 -----
Db 20916 TCGTCTGTGCTGTGCGCTGTGCATGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG 20975
QY 78 -----
Db 20976 GCTCTGTGCGCGTTCCTTCTCTCTCTGTGGTCTCTCTGTGCTCATCCCATCTCGCCCG 21035
QY 78 -----
Db 21036 CCCCATCCAGCCCTTCTCCCGGCTCCCACTGTGCCACACCTCCCGGCTCTCGCGCG 21095
QY 79 -ArgAlaLeuMetaspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlu 98
    |||||
Db 21096 CAGGGCGCTGATGAGACGAGCATGAAGAGATTGAAGGCTTCAAAATCGGAATCGGAGAGA 21155
QY 98 uGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAla 118
    |||||
Db 21156 ACAACTGACCCCGGTGGGGGAGAGACGCGGCGCTGTGCCAAGAGAGCTGCAGCGCGC 21215
QY 118 aglAlaArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGly 138
    |||||
Db 21216 GCAGGCGCGGTGGGGCGGACATGAGAGACTGTGGCGCGCTGTGTGCTGCTACCGCGG 21275
QY 138 yGluValGlnAlaMetLeuGlnSerThrGluGluLeuArgValArgLeuAlaSerHis 158
    |||||
Db 21276 CGAGGTGAGGCGATGTCTGTGGCGCAGAGCACCGAGAGAGCTGTGGGTGGCGCTTCGCCCA 21335

```



```
Db 396 ATGAGGTTCTGTGGGCTGCTGTGTCACATTCTTGAGGATGCCAGGCCAAGGTG 337
Qy 21 GtGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGlnTrpGlnSer 40
Db 336 GAGCAAGCGGTGGAGACACAGCCGGAGCCNCAGACTGCCAGCAGACCCAGTGGCAGAC 277
Qy 41 GtGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 276 GGCACAGCGTGGGAACCTGGCACTGGGTCCCTTTGGGATTACTGGCGCTGGGTGCAGACA 217
Qy 61 LeuSerGluGlnValGlnGlnGluLeuLeu-SerSerGlnValTrpGlnGluLeuArgAl 80
Db 216 CTGTCTGACACAGTGCAGAGAGAGCTGCTCGAGCTCCAGTCCACCAGCACTGAGGCGC 157
Qy 80 aLeuMetAspGluTrpMetLeuGlyGluLeuLysAlaTyrLysSerGluLeuGlnGlnLe 100
Db 156 GCGATGGACGAGACCATGAAGAGCTTGAAGGCTTACAAATCGAATCGAGAACCAACT 97
Qy 100 uThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAl 120
Db 96 GACCCCGGTGGGAGAGACCGGGCGACGCTGTCCAAAGAGCTGCAGCGCGCGCAGAGN 37
Qy 120 aArgLeuGlyAlaAspMetGluAspVal 129
Db 36 CCGGCTGGGCGGAGCATGAGAGACGTG 9

RESULT 15
US-09-880-107-2491/c
; Sequence 2491, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIORITY FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2491
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 N33009
; LOCATION: (1)..(478)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-2491

Alignment Scores:
Pred. No.: 2,69e-48 Length: 478
Score: 625.00 Matches: 126
Percent Similarity: 97.69% Conservative: 1
Best Local Similarity: 96.92% Mismatches: 2
Query Match: 39.53% Indels: 1
DB: 10 Gaps: 0

US-09-827-854-14 (1-317) x US-09-880-107-2491 (1-478)
Qy 1 MetLysValLeuTrpAlaLeuLeuValTrpPheLeuAlaGlyCysGlnAlaLysVal 20
Db 396 ATGAGGTTCTGTGGGCTGCTGTGTCACATTCTTGAGGATGCCAGGCCAAGGTG 337
Qy 21 GtGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGlnTrpGlnSer 40
Db 336 GAGCAAGCGGTGGAGACACAGCCGGAGCCNCAGACTGCCAGCAGACCCAGTGGCAGAC 277
Qy 41 GtGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 276 GGCACAGCGTGGGAACCTGGCACTGGGTCCCTTTGGGATTACTGGCGCTGGGTGCAGACA 217
Qy 61 LeuSerGluGlnValGlnGlnGluLeuLeu-SerSerGlnValTrpGlnGluLeuArgAl 80
Db 216 CTGTCTGACACAGTGCAGAGAGAGCTGCTCGAGCTCCAGTCCACCAGCACTGAGGCGC 157
Qy 80 aLeuMetAspGluTrpMetLeuGlyGluLeuLysAlaTyrLysSerGluLeuGlnGlnLe 100
Db 156 GCGATGGACGAGACCATGAAGAGCTTGAAGGCTTACAAATCGAATCGAGAACCAACT 97
Qy 100 uThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAl 120
Db 96 GACCCCGGTGGGAGAGACCGGGCGACGCTGTCCAAAGAGCTGCAGCGCGCGCAGAGN 37
Qy 120 aArgLeuGlyAlaAspMetGluAspVal 129
Db 36 CCGGCTGGGCGGAGCATGAGAGACGTG 9
```

Search completed: March 15, 2003, 03:03:28
Job time : 82.8753 secs

```
Qy 41 GtGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 276 GGCACAGCGTGGGAACCTGGCACTGGGTCCCTTTGGGATTACTGGCGCTGGGTGCAGACA 217
Qy 61 LeuSerGluGlnValGlnGlnGluLeuLeu-SerSerGlnValTrpGlnGluLeuArgAl 80
Db 216 CTGTCTGACACAGTGCAGAGAGAGCTGCTCGAGCTCCAGTCCACCAGCACTGAGGCGC 157
Qy 80 aLeuMetAspGluTrpMetLeuGlyGluLeuLysAlaTyrLysSerGluLeuGlnGlnLe 100
Db 156 GCGATGGACGAGACCATGAAGAGCTTGAAGGCTTACAAATCGAATCGAGAACCAACT 97
Qy 100 uThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAl 120
Db 96 GACCCCGGTGGGAGAGACCGGGCGACGCTGTCCAAAGAGCTGCAGCGCGCGCAGAGN 37
Qy 120 aArgLeuGlyAlaAspMetGluAspVal 129
Db 36 CCGGCTGGGCGGAGCATGAGAGACGTG 9
```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:09:47 ; Search time 1289.69 Seconds

(without alignments)
3980.771 Million cell updates/sec

Title: US-09-827-854-14

Perfect score: 1881

Sequence: 1 MKVLMALVTLFLAGCCQAKV.....VEKQAAVGTSAAPVPSDNH 317

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 1615406 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n_model -DEV=xlp

-O=/cgn2.1/USPTO.spool/US09827854/funat_11032003_101610_27495/app_query.fasta_1.3576

-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0

-UNITS-bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45

-DOCALLIGN=200 -NRM=ext -HEAPSIZE=500 -MINLEN=0 -ALIGN=15 -MODE=LOCAL

-USER=us09827854.cgn_1.1_8826_ernat_11032003_101610_27495 -NCPU=6 -ICPU=3

-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV=TIMEOUT=120

-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST :

1: em_estba :

2: em_esthum :

3: em_estlin :

4: em_estmu :

5: em_estov :

6: em_estpl :

7: em_estro :

8: em_hlc :

9: gb_estl :

10: gb_est2 :

11: gb_hlc :

12: gb_est3 :

13: gb_est4 :

14: gb_est5 :

15: em_estfun :

16: em_estom :

17: gb_gss :

18: em_gss_hum :

19: em_gss_inv :

20: em_gss_pln :

21: em_gss_vit :

22: em_gss_fun :

23: em_gss_mam :

24: em_gss_mus :

25: em_gss_other :

26: em_gss_pro :

27: em_gss_rod :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	1278.5	80.9	1027	13	B1670367
2	1242	78.6	842	12	BG763371
3	1236	78.2	933	13	B1668318
4	1208	76.4	938	12	BG761746
5	1205.5	76.2	942	13	B1600906
6	1191	75.3	800	13	BM042094
7	1180.5	75.3	927	12	BG472299
8	1187.5	75.1	922	13	B1597743
9	1178.5	74.5	817	12	BG774871
10	1177	74.4	811	13	B1600563
11	1163.5	73.6	845	12	BG829472
12	1140	72.1	790	12	BG707147
13	1137	71.9	919	13	B1551475
14	1136	71.9	706	14	BM728696
15	1134	71.7	1110	11	AK010261
16	1127	71.3	757	13	BM042228
17	1117.5	70.7	907	12	BG706129
18	1116	70.6	741	12	BG762924
19	1112	70.3	804	12	BG702752
20	1111	70.3	834	13	BM042676
21	1108.5	70.1	808	13	B1668329
22	1105	69.9	855	13	B161362
23	1102	69.7	803	13	B1670350
24	1099	69.5	782	12	BG716776
25	1099	69.5	912	13	B1601551
26	1099	69.5	865	14	BQ677266
27	1096	69.3	757	13	B1603658
28	1093.5	69.2	797	12	BG715366
29	1093	69.1	846	13	B1159757
30	1092	69.1	796	13	BM042153
31	1092	69.1	1100	14	BM914382
32	1087	68.8	757	12	BG707750
33	1087	68.8	794	13	B1601279
34	1087	68.8	954	12	BF967543
35	1085	68.6	790	13	B1551066
36	1084	68.6	790	13	B1551811
37	1080.5	68.3	798	12	BG708414
38	1079.5	68.3	891	13	B1549292
39	1073.5	67.9	802	13	B1458355
40	1071.5	67.8	914	13	B1603523
41	1071	67.7	748	13	B1553085
42	1070.5	67.7	930	12	BF967857
43	1066.5	67.5	812	12	BG769968
44	1066	67.4	688	13	B1544886
45	1065.5	67.4	803	12	BG709360

ALIGNMENTS

RESULT 1
LOCUS B1670367 1027 bp mRNA linear EST 12-SEP-2001
DEFINITION 603292738P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5312024 5',
ACCESSION B1670367 mRNA sequence.
VERSION B1670367
KEYWORDS B1670367.1 GI:15584600
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1027)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroaki Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLM11790 row: 1 column: 09
 High quality sequence stop: 845.
 Location/Qualifiers
 1. 1027
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:5312024"
 /tissue_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescript (modified pBluescript KS+); Site:1: BamHI; Site:2: SalI; XhoI (gtcgag)
): Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTAA-3', size selected for average insert size 2.3 kb and normalized to 10⁷ 5. This is a primary library enriched for full-length clones and constructed using the cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC library."
 BASE COUNT 194 a 308 c 400 g 125 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,79e-124 Length: 1027
 Score: 1278.50 Matches: 291
 Percent Similarity: 90.55% Conservative: 6
 Best Local Similarity: 88.72% Mismatches: 20
 Query Match: 80.87% Indels: 12
 Gaps: 2
 DB: 13
 US-09-827-854-14 (1-317) x B1670367 (1-1027)
 QY 1 MettysValLeutrrpAlaalaLeuLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 26 ATGAAGTCTCTGTGGCTGCTGTGCTGCATTCCTCGACAGATGCCAGGCCAAGTGT 85
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluGlnSer 40
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 86 GAGCAAGCGGTGAGACAGAGCCGAGCCGACCTGGCCAGACAGACCGAGTGGCAGAGC 145
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 146 GGGCAGCGCTGGGAGACGACGCGGCTTTGGGATTACTGCGCTGGGTGGCAGACA 205
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 206 CTCTCTGAGCAGCTGACGAGGAGCTGCTCAGCTCCACAGGTCAACGAGAACTGAGGGCG 265
 QY 81 LeuMetAspGluThrMetLeuGluLeuValThrLysSerGluLeuGlnAlaAlaGlnAla 100
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 266 CTGATGAGCAGACCATGAAGAGGTTGAGGCTTACAAATCGGACACTGGGAGAACACTG 325
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 326 ACCCGGTTGGCGGAGAGACGGCGGACGCTTCAAGAGCTGACGAGGGCGGCGAGGCC 385
 QY 121 ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnThrArgGlyGluVal 140
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 386 CGGCTGGGGCGGACATGTGAGGACGTGTGGCGCGCTGTGACATACCGCGCGAGAGTGT 445
 QY 141 GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 446 CAGGCCATGCTGGCCAGACAGACCAGAGAGCTCGGGTGGCTTCCTCCACCTGGCCG 505
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 506 AACCTGGGTAAAGGGGCTCTCCCGATGCCGATGACCTGACAGAGCGCTGGCAGTGTAC 565
 QY 181 GlnAlaGlyAlaArgGluGluGluAlaGluArgGlyLeuSerAlaIleArgGluArgLeu-Gl 200
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 566 CAGGCGGGGCGCGCGAGCGCGCGAGCGCGGCTCAGCCATCCGCGAGCGCTGGGG 625
 QY 200 YProLeuValGluGlnGluArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnR 220
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 626 CCCCTGGTGGAAACAGGTCCGCTGGCGGCGCGCTGCTCCCTGTGGCCGACGCC 665
 QY 220 OleuGlnGluArgAlaGlnAlaThrPglGluArg-LeuArgAlaArgMetGluGlu-Met 239
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 686 GCTACAGAGACGGGCCAAGGCTGGGGCCAGCGGCTTGGCGCGCGGATGGAGAGAGAG 745
 QY 240 GlySerArg-ThrArgAspArgLeuAspGluValLysGluGlnValAlaGlu-ValArgA 259
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 746 GGCAGCGCGGAGCCCGGACCGCTGACCGAGGTGAAGAGACAGGTGGCGGAGGTGCGCG 805
 QY 259 LAlsLeu---GluGluGlnAlaGlnGlu-IleArgLeuGln-AlaGluAlaPheGlnAl 277
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 806 CCCAAGCTGACGAGCGAGCCAGCCAGACAGATACGCTTGAGGCGCGAGGCTTCTAGGC 865
 QY 277 ArgLeuLysSerTrpPheGluPro-----LeuValGluAspMetGlnArgGlnTrpAl 295
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 866 CCGGCTCAAGAGCTGGTTCACACCCCTGGCTTGAACAGACATAGGCGCGGCCACAGTGGGG 925
 QY 295 AGIleuValGluLysVal-GlnAlaAlaValGlyThrSerAlaAlaProValProSerA 315
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 926 CGGCTGTGGAAAAGGTGTCAGGCTCGCGGGGACACACCGCGCTGTGTCCAGC-G 984
 QY 315 spasnHis 317
 ||||||||
 DB 985 ACAATCAT 992
 RESULT 2
 BG763371 842 bp mRNA linear EST 15-MAY-2001
 LOCUS 602735433F1 NIH_MGC_49 Homo sapiens CDNA clone IMAGE:4860585 5',
 DEFINITION mRNA sequence.
 ACCESSION BG763371
 VERSION BG763371.1 GI:14074024
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 842)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabs-r@mail.nih.gov
 Tissue Procurement: ATCC/DC/DTP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLM1718 row: b column: 10
 High quality sequence stop: 817.
 Location/Qualifiers
 1. 842
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:4860585"
 /clone_lib="NIH_MGC_49"
 /tissue_type="melanotic melanoma, high MDR (cell line)"
 /lab_host="DH10B (phage-resistant)"

Db	195	GGCCACGCTGGGAACCTGCACACTGGGTCCCTTTGGATTACCTGCGGTGGGTCAGACA	254
Qy	61	LeuSerGIuGIuValGIuGIuGIuLeuLeuSerSerGIuValThrInGIuLeuAArgAla	80
Db	255	CTGTGTGACAGAGTGCAGAGAGAGCTGTACAGTCTCCAGCTCACCCAGAACTGAGAGGGG	314
Qy	81	LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGIuGIuLeu	100
Db	315	CTGATGACAGACACCTTGAAGAAGTTGAAGGCTTACAAATTCGAACTGAGACACACTG	374
Qy	101	ThrProValAlaGIuGIuGIuThrArgAlaArgLeuSerLysGIuLeuGIuAlaAlaGIuAla	120
Db	375	ACCCCGGTGGCGAGAGACGCGGGCACGCTGTCTCCAAAGAGCTGCAGGCGGGCGAGGCC	434
Qy	121	ArgLeuGIuAlaAspMetGIuAspValArgGIuArgLeuValGIuTyrArgGIuVal	140
Db	435	CGGTGTGGCGCGACATGTGAGAGAGCTGTCCGGCCGCTGTGTGCATGACCGCGCGAGGTG	494
Qy	141	GIuAlaMetLeuGIuLysInSerThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg	160
Db	495	CAGGCGATGCTCGGCCACAGACACCGAGAGCTCGGGTGGCTGCTCCACCTCGCCG	554
Qy	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGIuLysArgLeuAlaValTyr	180
Db	555	AAGCTCGTAAAGCGCTCTCCGCGATGCGCATGACCTGCAGAAAGCGCTGGGAGTGC	614
Qy	181	GIuAlaGIuAlaArgGIuGIuGIuAlaArgGIuLysSerAlaAlaArgGIuArgLeuGIu	200
Db	615	CAGGCGGGGGCGCGCAGAGGCGCGCGAGCGGCTTCAGGCCATCGCAGAGCGGCTGGGG	674
Qy	201	ProLeuValGIuGIuGIuLysArgValArgAlaAlaThrValGIuSerLeuAlaGIuInPro	220
Db	675	CCCCGTGTGGAACAGGAGCGCGCTGCGGGCGGCCACTGTGGCTCCTCGGCGGACAGCGG	734
Qy	221	LeuGIuGIuArgAlaGIuAlaThrArgGIuGIuArgLeuArgAlaArg--GluLeu	239
Db	735	CTACAGAGACGGGGCCAGGCTGTGGGCGAGCGGCTGCGCGCGATTTGGAGGCACATTG	794
Qy	239	tGlySerArgThrArgAspArgLeuAsp--GluValLysGIuGIuValAlaGIuValArgA	259
Db	795	GGGCAAGCCCGAGCCCGGAGGCTGTGAGAGAGTGAAGAGCAGGTGCGGAATTGGCGCC	854
Qy	259	1aLysLeuGIuGIuGIuGIuAlaGIuGIuLysArgLeuGIuAlaGIuLysArgAlaArgL	279
Db	855	GCCAAAGCTGGGGGAGCACAGGCCCGCAATTGCTTGCAGGCCAA--GCCTTCCAGGCGGCT	913
Qy	279	eulLysSerTrpPheGIu	284
Db	914	CCAAAGAGCTGTGTTCAA	930
RESULT 4			
LOCUS	BG761746	938 bp	mRNA linear EST 15-MAY-2001
DEFINITION	602171742P1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4841411 5',		
ACCESSION	BG761746		
VERSION	BG761746.1	GI:14072399	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	NIH-MGC http://mgi.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabs-remail.nih.gov		
	Tissue Procurement: ATCC/DC/DTF		
	cDNA Library Preparation: Lind Hong/Rubin Laboratory		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		

found through the I.M.A.G.E. Consortium/LLNL at
http://image.llnl.gov
Plate: ILCM1674 row: c column: 12
High quality sequence stop: 767.

```

source
1. .938
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4841411"
/clone_1ib="NIH_MGC_49"
/lusue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCCGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

```

Pred. No.:	8 72e-117	length:	938
Score:	1208.00	Matches:	271
Percent Similarity:	92.28%	Conservative:	4
Best Local Similarity:	90.94%	Mismatches:	12
Query Match:	76.94%	Indels:	11
DB:	12	Gaps:	2

QY	1	MettysValLeuThrPalaIalaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	57	ATGAAGGTTCTGGGGCTCGCTGGTGGTCAATCTCCGAGAGATGCCAGGCCAAGGGT	11
QY	21	GIuGlnAlaValGIuThrGIuProGIuLeuArgGlnGlnThrGIuTrpGlnSer	40
Db	117	GAGCAAGCGGTGAGACAGACCGGAGGCCGAGCTGGCGCCAGCAGACCGATGGCAGAC	17
QY	41	GIyGlnArgTrpGluLeuAlaLeuGlnLysArgPheTrpAspTrpLeuArgTrpValGlnThr	60
Db	177	GCGCAGCGCTGGGAACCTGGACCTGGGTGCTTTGGGATTACTCGCGCTGGGTGCAGACA	23
QY	61	LeuSerGIuGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80
Db	237	CTGTCTGACGAGTGCAGAGAGAGCTGTCAAGTCCAGATCCACGAGAACTGAGGGCG	29
QY	81	LeuMetAspGIuThrMetLysGIuLeuLysAlaTrpLysSerGIuLeuGlnGlnGlnLeu	100
Db	297	CTGATGGACGAGACCATGAAGAGTTGAAGGCTACAAATCGGAACTGGAGGAACTG	35
QY	101	ThrProValAlaGluGluThrArgAlaArgLeuSerLysGIuLeuGlnAlaIalaGlnAla	120
Db	357	ACCCCGGTGGCGAGAGACGGGGCAAGCTGTCAAGAGAGCTGCAGGCGGCGCAGGCC	41
QY	121	ArgLeuGlnValAspMetLysPheValArgLysArgLeuValGlnThrArgGlyGluVal	140
Db	417	CGGCTGGCGCGGACATGGAGAGACTGTGGCGGCCCTGGTGCATGACCGCGGCGAGGTG	47
QY	141	GlnAlaMetLeuGlnGlnSerThrGIuGluLeuLeuArgValArgLeuAlaSerHisLeuArg	160
Db	477	CAGGCCATGCTGGCCAGACGACGAGSAGCTGGGGGTGGCGCTGCCCTCCACCTGGGC	53
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspSerLeuGlnLysArgLeuAlaValTrp	180
Db	537	AAAGCTCGTAAGCGGCTCTCCGCGATGCCATGACTCGCAGAAAGCGCTGGCAGCTGAC	59
QY	181	GlnAlaGlyAlaArgGlnGlyAlaGlnArgLysLeuSerAlaIleArgGluArgLeuGly	200
Db	597	CAGGCGGGAGCCCGGAGGGCGCCAGCGCTGTGACGCCCATTCGCGAGGCGCTGGGGG	65

Best Local Similarity:	91.07%	Mismatches:	14
Query Match:	76.25%	Indels:	8
DB:	13	Gaps:	1
US-09-827-854-14 (1-317) x B1600906 (1-942)			
QY 1	MettysValLeuTPRAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20		
DB 73	ATGAAGGTTCTGTGGGCTCGCTGCTGGTCAATTCCTGGCAGATGCCAGGCCAAGG 132		
QY 21	GlutGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40		
DB 133	GAGCAAGCGGTGACACAGAGCGGAGCCCGAGCTGCGCAGACACAGCAGAGTGGCAGAGC 192		
QY 41	GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60		
DB 193	GGCCAGCCCTGGGAACTGGCACTGGTGGCTTTGGATTACCTCGCGCTGGGTGCAGACA 252		
QY 61	LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80		
DB 253	CTGTGTGAGCAGTCCAGAGAGAGCTGCTCACCTCCAGGTCCACCAAGAACGTGAGGCG 312		
QY 81	LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGlnLeuGlnGluGlnLeu 100		
DB 313	CTGAATGAGACGAACCATGAAGAGTTGAAGGCTTACAAATCGGAATCGAGAGAACACATG 372		
QY 101	ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120		
DB 373	ACCCGGTGGCGGAGAGACAGCGGGCAGCGGTCTCCAAAGAGCTCAGCGGCGCAGGCG 432		
QY 121	ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTrpArgGlyGluVal 140		
DB 433	CGGCTGGGCGCGGACATGAGAGACTGTGCGGCGCCCTGATGTCACATACCGCGCAGAGTG 492		
QY 141	GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160		
DB 493	CAGGCCATGCTCGGCGACAGACACCGAGAGCTGCGGGTGCCTGCTGCCACCTGCGGC 552		
QY 161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180		
DB 553	AAGCTGCGTAAGCGGCTCTCCGCGATGCGCATGTACCTGCACAGACGCGCTGGCAGTATAC 612		
QY 181	GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200		
DB 613	CAGCGCGGGGCGCGGAGGCGCGCAGCGCGGCGCTCACAGCGCATCCGCGAGCGCTGGGG 672		
QY 201	ProLeuValGluGlnGlyArgValArgAlaIleThrValGlySerLeuAlaGlyGlnPro 220		
DB 673	CCCTGGTGGAAACAAGGCGCGCTGCGGGCG-GCCACTGTGGGTCTCTGGCGGCGACGCG 731		
QY 221	LeuGlnGluArgAlaGlnAlaArgPheGlyLeuArgLeuArgAlaArgMetGlu-GluMetG 240		
DB 732	CTACAGAGAGCGGCGCCCAAGCTCGGGGCGACAGCGGTCGCGCGCGGATGTGAGAGCAGATTG 791		
QY 240	LysSerArgThrArgAspArgLeuAsp-GluValLys-GluGlnValAlaGluValArg 259		
DB 792	GCAGCGGAGACCGCGCAGCGGCTGAGCAGGAGGAGACAGACAGGTTGGGAAATTGGCGG 851		
QY 259	IaLysLeuGluGluGlnAla---GlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 278		
DB 852	CCAACCTGGAGAGAACAGCGCCACGACGAGATACCTCTTGACAGGCGCAGGCGCTTCAGGCGCC 911		
QY 278	rgLeuLysSerTrpPheGluPro 285		
DB 912	GCCTCAAAAC-TGGTTGACCCCT 933		

RESULT 6

BM042094

LOCUS

DEFINITION

ACCESSION

VERSION

BM042094

603615713p1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5420618 5', mRNA sequence.

BM042094

BM042094.1 GI:16771361

800 bp

mRNA

linear

EST 07-NOV-2001

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 800)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTP/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LNCMI875 row: 1 column: 03
High quality sequence stop: 792.

FEATURES
source
1..800
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5420618"
/clone_lib="NIH-MGC_112"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site: 1; XhoI: Site 2; EcoRI: cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 146 a 239 c 315 g 100 t
ORIGIN

Alignment Scores:
Pred. No.: 4.28e-115 Length: 800
Score: 1191.00 Matches: 243
Percent Similarity: 98.80% Conservative: 4
Best Local Similarity: 97.20% Mismatches: 3
Query Match: 75.33% Gaps: 1
DB: 13 Gaps: 0

US-09-827-854-14 (1-317) x BM042094 (1-800)

QY 1 MetTysValLeuTrpAlaAlaLeuValTrpPheLeuAlaGlyCysGlnAlaTysVal 20
|||||
DB 50 ATGAAGGTTCTGTGGGCTGGTGTGTCACATTCCTGCGAGAGATGCCAGGCAAGTG 109
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
|||||
DB 110 GAGCAAGCGGTGAGACAGACAGCGGAGCCGAGCTCCGACACAGACCCAGATGGCGAGAC 169
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
DB 170 GGGCAGCGCTGGAACTGGCAGCTGGCTTTTGGGATTACCTGGCTGGGTGGAGATA 229
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
|||||
DB 230 CTGTCTGACAGAGTGCAGAGAGAGTGTACAGCTCCAGAGTCAACCCAGAACTGAGAGCGG 289
QY 81 LeuMetAspGluThrMetTysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
|||||
DB 290 CTGAATGAGACCACTGAAGAGATTGAAGGCTTAAATCGAATCGAGAAACAACATG 349
QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
|||||
DB 350 ACCCCGGTGGCGAGAGACGCGGCGACGCTGTCTCAAGAGACTGAGCGGCGAGGCC 409

QY 121 ArgLeuGlyAlaAspMetClnuAspValArgGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
DB 410 CGGCTGGCGCGGACATGAGAGACGTGTGCGGCGGCTGTCAGTACCGCGGAGGTG 469
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
DB 470 CAGGCCATCTCGGCGAGAGACCGAGAGAGCTCGGGGTGGCTCCCTCCACCTGGCCG 529
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
DB 530 AAGCTCGTAAGGGGCTCTCCGCGATGCGATGACCTCAGAAAGCGCTGGCAGGTAC 589
QY 181 GlnAlaGlyAlaArgGluGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||
DB 590 CAGGCGGGGCGCGCAGAGGCGCGGAGCGGCTCAGCGCCATCCGACGCGCTGGGGG 649
QY 201 ProLeuValGluGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
|||||
DB 650 CCCCTGTGTGAACAGGCGCGCGTGGCGGACAGCTGTGGGCTCCCTGGCGGCGGCGG 709
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
|||||
DB 710 CTACAGAGACGCGGCGCCAGGCTGTGGGCGAGCGGTG-CGCGCGCGGATGAGAGATGGGC 768
QY 241 SerArgThrArgAspArgLeuAspGluVal 250
|||||
DB 769 ACGCGGAGCGCGGACGCTGTGAGCAGCGTG 798

RESULT 7
BG472299
LOCUS BG472299 927 bp mRNA linear EST 21-MAR-2001
DEFINITION 602513830F1 NIH-MGC_16 Homo sapiens cDNA clone IMAGE:4645759 5',
mRNA sequence.
ACCESSION BG472299
VERSION BG472299.1 GI:13404485
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 927)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LNCMI419 row: k column: 08
High quality sequence stop: 848.

FEATURES
source
1..927
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4645759"
/clone_lib="NIH-MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pORF7; Site: 1; XhoI: Site 2; EcoRI: cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 170 a 271 c 371 g 115 t
ORIGIN

Alignment Scores:

Pred. No.: 5,95e-115 Length: 927
 Score: 1190.50 Matches: 274
 Percent Similarity: 92.00% Conservative: 2
 Best Local Similarity: 91.33% Mismatches: 14
 Query Match: 75.30% Indels: 12
 DB: 12 Gaps: 1

US-09-827-854-14 (1-317) x BG472299 (1-927)

OY 1 MettysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaIalysVal 20
 |||||
 DB 47 ATGAAGGTTCTGTGGCTGCTGTGTCATTCCTCGCAGATGCCAGGCAAGTGC 106
 |||||
 OY 21 GluGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 |||||
 DB 107 GAGCAAGCGGTGAGACAGACCGGAGCCGAGCTCCGACACAGACGAGGCGAGAC 166
 |||||
 OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 |||||
 DB 167 GGCCAGCGCTGGGAAGTGGCACTGGCTGCTTTGGGATTACTGCGCTGGGTCAGACA 226
 |||||
 OY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||||
 DB 227 CTGTCTGACAGGTGAGAGAGAGCTGCTCAGCTCCAGGTACCCAGGAACTGAGGGCG 286
 |||||
 OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGluLeu 100
 |||||
 DB 287 CTGATGAGACGACACCTGAGAGAGTTCGAGGCTTCACAAATCGGAAGTGAAGACAACTG 346
 |||||
 OY 101 ThrProValAlaGluGluThrArgAlaLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 |||||
 DB 347 ACCCGGCTGGGAGAGAGCGGGGACGCTGCCAAGAGACTGAGGCGGCGAGGCG 406
 |||||
 OY 121 ArgLeuGlyAlaAspMetCysPylArgGlyArgLeuValGlnTrpArgGlyGluVal 140
 |||||
 DB 407 CGGCTGGGCGGACATGAGAGACGCTGCGCGCTGCTGTCAGTACCGCGGCGAGGTG 466
 |||||
 OY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||
 DB 467 CAGGCCATCTCTGGCCAGAGACCCGAGAGCTGCGGCTGCGCTGCCCTCCACCTGCGCG 526
 |||||
 OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||||
 DB 527 AAGCTGCGTAAGCGGCTCTCCGCGATGCCGATGACCTCAAGAGCCCTGGCAGGTAC 586
 |||||
 OY 181 Gln-AlaGlyAlaArgGlyAlaGluArgGly--LeuSerAlaIleArgGluArgLeu 199
 |||||
 DB 587 CAGGGCGGGGCGCGGAGAGCGCGGAGCGGCGCTCCAAAGCCCATCCGCGAGCGGCTT 646
 |||||
 OY 200 --GlyProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGly 219
 |||||
 DB 647 GGGGCGGCTGGGTGGACAGGCGCGGCTGCGGCGGCGCCTGTGGCTGCGCGGCGG 706
 |||||
 OY 219 LysProLeuGlnLysArgAlaGlnAlaTrpGlyLysArgLeuArgAlaArgMetGluGln 239
 |||||
 DB 707 AAGCG-CTACAGAGAGCGGCGCGGCTGGGCGGAAGCGTGC-CGCGGCGGAGTGAAGANA 764
 |||||
 OY 239 etGlySerArg-ThrArgAspArgLeuAspGluValLysGluGlnValaGlu-ValArg 258
 |||||
 DB 765 TGGGGAGAGGAGACCGCGAGCTGAGCCGAGGTGAAGAGACAGTGGCGAGGGGTGG 824
 |||||
 OY 258 gAlaLysLeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAla---PheGlnAl 277
 |||||
 DB 825 CCGCAAGCTGGAGGAAGCAGGCGCAGCAGATACGCTGCAGGCGGAGGGGCTTCCAGGGC 884
 |||||
 OY 277 aArg-LeuLysSerTrpPhe-GluProLeuValGluAspMet 290
 |||||
 DB 885 CCGGCTCAAGAGAGTGTTCGAGGCCCTGTGGTGAAGACATG 926
 |||||
 RESULT 8
 B1597743 LOCUS B1597743 922 bp mRNA linear EST 07-sep-2001

DEFINITION 603248609P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5300259 5',
 mRNA sequence.
 ACCESSION B1597743
 VERSION B1597743.1 GI:15490682
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 922)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE JOURNAL
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
 Plate: LLM11760 row: b column: 04
 High quality sequence stop: 782.
 FEATURES
 source
 1..922
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5300259"
 /clone_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: Brain; Vector: pBluescript (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.3 kb and
 normalized to R0F 5. This is a primary library enriched
 for full-length clones and constructed using the
 CAP-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH-MGC library."
 BASE COUNT 176 a 273 c 359 g 114 t
 ORIGIN

Alignment Scores:

Pred. No.: 1.22e-114 Length: 922
 Score: 1187.50 Matches: 256
 Percent Similarity: 93.19% Conservative: 4
 Best Local Similarity: 91.76% Mismatches: 15
 Query Match: 75.11% Indels: 5
 DB: 13 Gaps: 1

US-09-827-854-14 (1-317) x B1597743 (1-922)

OY 1 MettysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaIalysVal 20
 |||||
 DB 73 ATGAAGGTTCTGTGGCTGCTGTGTCATTCCTCGCAGATGCCAGGCAAGTGC 132
 |||||
 OY 21 GluGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 |||||
 DB 133 GAGCAAGCGGTGAGACAGAGCCGAGCGGAGCTGCCAGACAGACGAGGCGAGAGC 192
 |||||
 OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 |||||
 DB 193 GGCCAGCGCTGGGAAGTGGCACTGGCTGCTTTGGGATTACTGCTGCTGGCTGAGACA 252
 |||||
 OY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||||
 DB 253 CTGTCTGACAGGTGACAGAGAGCTGCTCAGCTCCAGGTACCCAGGAACTGAGGGCG 312
 |||||
 OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGluLeu 100
 |||||

Db	313	CTGATGGACGAGACATGTAAGAGATTGAAGGCTTCACAAATGGCAACTGGAGGAACAACACTG	372		
OY	101	ThrProValAlaGluGluThrArgAlaArgSerLysGluGlnAlaAlaGlnAla	120		
Db	373	ACCCGCGTGGCGGAGACACGGCGGACGGCTGTCTCAAGGAGCTGCAGCGCGGAC	432		
OY	121	ArgLeuGlyAlaAspMetGlnAspAlaArgGlyArgLeuValGlnTyrArgGlyVal	140		
Db	433	CGGCTGGGCGCGAATGAGAGACCTGTGGCGGCCCTGTGTCACTACCGGCGGAGGTG	492		
OY	141	GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg	160		
Db	493	CAGGCCATGCTGGCGGACAGCACCGAGGAGCTGGCGGTGGCGCTGCCCTCCACTGGGC	552		
OY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180		
Db	553	AAGCGCTTAACGGGCTCTCTCCGATCCCGATGACCTGCAGAACCGCTGGCATGTAC	612		
OY	181	GlnAlaGlyAlaArgGluGluValaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200		
Db	613	CAGGCCGGGGCCCCGGAGGGCGCCGAGGGCGGCTC-AGCGCCATCCGGAGCGCTGGGG	671		
OY	201	ProLeuVal-GluGlnGlyArgValaArgAlaAlaThrValGlySerLeuAlaGlyGln-P	220		
Db	672	CCCCGTGGTGGAGAACAGGGCGGGGTCCGGCGCCCACTGTGGCTCTCGCGCGGACCC	731		
OY	220	LeuLeuGluGluArg-AlaGlnAlaThrPglGlyGluArgLeuArgAlaArgMetGluMet	239		
Db	732	CGCTCAGAGAGCGGGGCCAGGCTCTGGGGCGAGCGGATGACGCGCATGTGACGAGCAGG	791		
OY	240	GlySerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValaArgAla	259		
Db	792	GGCACCAGGACCCGAGACCGCTGTGACGAGGTAAGAGAGACAGTGGCGGGGTGGGCCA	851		
OY	260	LysLeuGluGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGln	276		
Db	852	CAAGCTGAAGAAACAGCCCAAGAAATACGCTTA---GCCGAAGGCTTCCAGG	899		
RESULT 9	BG774871	817 bp	mRNA	linear	EST 15-MAY-2001
LOCUS	BG774871				
DEFINITION	602645975P1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4760827 5',				
ACCESSION	BG774871				
VERSION	BG774871.1	GI:14045188			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	NIH-MGC http://mhc.nci.nih.gov/ .				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
COMMENT	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgapbs-remail.nih.gov				
	Tissue Procurement: DCTD/DTP				
	cDNA Library Preparation: Ling Hong/Rubin Laboratory				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)				
	DNA sequencing by: Incyte Genomics, Inc.				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LNL at:				
	http://image.lnl.gov				
	Plate: L10CM1612	row: e	column: 20		
	High quality sequence stop: 813.				

```
location/Qualifiers
1. .817
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4760827"
/clone_lib="NIH.MGC.40"
/rtsuse_type="carchina, cell line"
/lab_host="DH10 (phage-resistant)"
```

```

/Note="Organ: prostate; Vector: POT87; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAC(G). Library constructed by
Ling Hong in the Laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)".
Note: this is a NIH-MGC Library."
BASE COUNT      148 a      242 g      324 g      103 t
ORIGIN
Alignment Scores:
Pred. No.:      9.12e-114      Length:      817
Score:          1178.50      Matches:      251
Percent Similarity: 96.93%      Conservative: 2
Best Local Similarity: 96.17%      Mismatches: 6
Query Match:      74.54%      Indels:      0
DB:              12      Gaps:      0

```

	US-09-827-854 -14 (1-317) x BG77487J(1-817)
OY	1 MettysValleutTPAlaIalaLeuleuValThrheLeuaIagIySGInalalyVal 20
Dd	40 ATGAAGGTTCTGTGGCGCTGCCTGTGTCAATTCCTGGCAGATGCCAGGCCAAGTG 99
OY	21 GluGlnAlaValGIuThrGlProGluProGluLeuArgGInGInThrTrpGInSer 40
Dd	100 GAGCAGAGGGTGGACAGACC66AGCC6AGCTGGCCAGCACAGACCAGACTGGCAAGC 155
OY	41 GlyGlnArqTPRGluLeuAbLaLeuGLIYARePhetRpaSPYrLeuArqTPVaIGInThr 60
Dd	160 GGCCAGCGCCTGGAAACTGGCACTGGGTGGCTTTGGATTACCTGGCTGGGTGCACACA 219
OY	61 LeuSerGIuGInValGInGluGluLeuLeuSerSercInValThrgInGluLeuArgAla 80
Dd	220 CTGTCTGAGCGAGGTGCAGAGAGACTGCTGCATGCCAGGTCAACCAGAACTGAGGGCG 279
OY	81 LeuueTaSPGIuThrMetLeYSgluLeuLySaIatyrLysSercIulaeugInGluGInLeu 100
Dd	280 CTGATGGACGAGACCATTAAGAGATTGAAGGCTTACAAATTCGAGACTGGAGAACACTG 339
OY	101 ThrProValaIaGluGluThrArgAlaIarqLeuSerTySGluLeuGlnAlaIalaGlnAla 120
Dd	340 ACCCGGGTGGGAGAGAACCCGGGACAGCGTCTCCAAGGAGCTGCAGCGGGCCAGGGCC 399
OY	121 ArgLeuGluIaIaAspmecIuaSpValAargLIyrrgIeuValGIntyrrargIyGluVal 140
Dd	400 CGGCTGGGGCGGCAATGAGAGAGCTGGCGGGCCCTGGTGGCAgyrAcCGGGCGAGGTG 455
OY	141 GlnAlamecLeuGlyGInSerThCIuGluLeuArqValArlgeLuAlaserHisLeuArg 160
Dd	460 CAGGCGATGCTCGGCCAAGACCGAGAGAGCTGGGGGTGGCCGTCCACTGCGCTCCACTGCGC 519
OY	161 LysLeuArqLySarqLeuLeuArqAspAlaAspAspleuGInLySarqLeuAlaValTyr 180
Dd	520 AAGCTCGCTAAGCGGCTCTCCGGATGCCATACCTGCAGAAAGCCCTGCGAGCTGTAC 579
OY	181 GlnAlaGlyAlaArqGluGlyAlaGluArgGlyLeuSerAlaIleArqGluArgLeuGly 200
Dd	580 CAGCGCGGGGGCCCGAGAGGGCGCGAGCGGCTCTCAAGCCATCCCGAGCGGCTGGGG 639
OY	201 ProLeuValGIuGInGlyArqValArlgeLuAlaIthrValGIyseIteLuAlaGlyGInPro 220
Dd	640 CCCCTGGTGGAAAGAGGGCGGGCTGGCGGCTGTGGGCTCTCTGGCCGGAG-CCG 699
OY	221 LeuGInGluArqAlaGlnAla--TRPGLyIuArqLeuArqAlaArqMetGluGInuMetG 240
Dd	699 CTACAGGAGCGGGCCAGGCCCTTGGGGCGAGCGATGGCCGGCCGATGAGAGAGATGG 758
OY	240 LyserArqThArqAspArqLeuAspGluValLySGluGInValAlaGluValArlgeAla 259
Dd	759 GCCACCGGACCG--GAAACCGCTTGACAGCGGTGAAGAGAGACAGTGGCGAGGTGGCGCC 815

RESULT 10
LOCUS B1600563
DEFINITION 603244936F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5287329 5',
mRNA sequence.
ACCESSION B1600563
VERSION B1600563.1 GI:15493502
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 811)
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM11726 row: 9 column: 10
High quality sequence stop: 783.
location/Qualifiers
1..811
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5287329"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pluescript (modified
pbluescript KS+); Site:1: BamHI; Site:2: SalI-XhoI (gtcagag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 148 a 245 c 316 g 102 t
ORIGIN
Alignment Scores:
Pred. No.: 1,3e-113 length: 811
Score: 1177.00 Matches: 240
Percent Similarity: 97.97% Conservative: 1
Best Local Similarity: 97.56% Mismatches: 4
Query Match: 74.45% Indels: 1
DB: 13 Gaps: 0
US-09-827-854-14 (1-317) x B1600563 (1-811)
QY 1 Mettysvalleutrpalaalaleulevalthrphelualaglycysglinalalysval 20
|||||
DB 74 ATGAGAGTTCTGTGTGGCGCTTCTGCTGACATTCCTGGCAGATGCCAGGCCAAGGTG 133
|||||
QY 21 GUGGlnalavalGluThrGluProGluProGluLeuArgGlnGlnThrGluUtrpGlnser 40
|||||
DB 134 GAGCAAGCGGTGAGACAGAGCCGAGCGCGCCGACACAGACGAGGAGGAGAGC 193
|||||
QY 41 GUGlnArqTrpGluLeuAlaLeuGlyArqPheTrpAspTrpLeuArgTrpValGlnThr 60
|||||
DB 194 GCGCAGCGCTGGAGACATGGGTGCGCTTTGGGATACCTCGCTGGTGGTACAGACA 253
|||||
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
|||||
DB 254 CTGCTGAGCAGAGTGCAGAGAGCTGCTCAGCTCCAGAGTACCCAGGAACACTGAGGCG 313
|||||

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu 100
|||||
DB 314 CTGATGAGACGAGACCATGAGAGGCTTGAAGCTCAATCGGAAGCTGAGGAACACTG 373
|||||
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
|||||
DB 374 ACCCGGTGGCGGAGAGACGCGGCGACGGCTGTCCAGAGACTGACGCGGCGAGGCC 433
|||||
QY 121 ArgLeuGlnAlaAspMetGluAspValArgGlyArgLeuValGlnTrpArgGlyGlnVal 140
|||||
DB 434 CGGCTGGCGCGGACATGAGAGACGCTGCGGCCCTGTGTCTACATACCGCGGAGGTG 493
|||||
QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
DB 494 CAGGCGATGCTGGCGCAGAGCACGAGAGAGCTGCGGCTGCGCTGCTCCCACTGGCG 553
|||||
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
DB 554 AAGCTGCGTAAGCGGCTCTCGCGCATGCGCATGACAGAACGCGCTGGCACTGATAC 613
|||||
QY 181 GlnAlaGlyAlaArgGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeu-Gl 200
|||||
DB 614 CAGGCGGCGGCGCGCGAGGCGCGCAGCGCGGCTCAGCGCCATCCGCGAGCGCTGGGG 673
|||||
QY 200 YProLeuValGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPr 220
|||||
DB 674 CCCCTGTGGTGAACAGGCGCGGCTCGGCGCCCACTGTGGCTCTCTGCGCAGCAGCC 733
|||||
QY 220 OLeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGlnMetGl 240
|||||
DB 734 GCTACAGAGCGCGGCGCCAGGCGCTGGCGAGCGGCTGGCGCGCGGAGTGAAGCAATGGG 793
|||||
QY 240 YSerArgThrArgAsp 245
|||||
DB 794 CAGCGGAAACCGCGAA 809
|||||
RESULT 11
LOCUS BG829472 845 bp mRNA linear EST 22-MAY-2001
DEFINITION 602763768F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4899112 5',
mRNA sequence.
ACCESSION BG829472
VERSION BG829472.1 GI:14177059
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 845)
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM1790 row: 9 column: 17
High quality sequence stop: 829.
location/Qualifiers
1..845
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4899112"
/clone_lib="NIH_MGC_42"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOT87; Site:1: XhoI;

OY	61	LeusertgUgInVaIGlGngUleuSerSergInValThrgInGUleuaArgAla	80
Db	253	CTGTCTGAACAGTCCAGAGAGACTGCTCAGCTCCAGGTACCAGAACTGAGGGCG	312
OY	81	LeumetAspGLutHrMetelySGluLeuysAlATyLySSerGlnUeUGluInUeu	100
Db	313	CtGAtGGAACGAGACATGAGAGATTGAAGCGCTTCMAATCGAAGCTGAGAGAACACTG	372
OY	101	ThrProValAlaGlUGluInThrArGAlaArgrLeuSerLySGluLeuGlnAlaIacInAla	120
Db	373	ACCCGGTGggCGAGAGAACGCCGGGCACGGCTGTCCAAGAGACTCGAGGGCGCAGGGC	432
OY	121	ArgLeuGlyAlaAspMecGluaspAlaArgGlyArgrLeuValGlnTyArGlyglUval	140
Db	433	CGGGTGGCGCGGACATGAGACGTGTGGCGCCCGCTGTGTCACTACCGCGGCGAGGTG	492
OY	141	GlnAlaMetLeuGlyGlnSerThrgUleuArGValArGrLeuAlASerHisLeuArg	160
Db	433	CAGGCATATGCTCGGCCAAGCACCCAGAGAGCTGGCGGTGCGCTCGCTCCACCTGGCGC	552
OY	161	LysLeuArgLySarGrLeuLeuArGrSpAlaasPaspLeuGlnLySarGrLeuAlaValTyr	180
Db	553	AAGCTGCGTAAGCGGCTCTCCGCACTCCGATGACCTGCAGAAGCGCTGGCACTGTAC	612
OY	181	GlnAlaGlyAlaArGrUGluGlyAlaGlnArgrGlyLeuSerAlaIlleArGrLuarGrLeuGly	200
Db	613	CAGGCGGGGGCCCCGAGAGCGGCCAGCGCGGTCTCACAGCCCATCGCGAGCGCTGGGG	672
OY	201	ProLeuValGlnUGlnGlyArGrValArGrAlaAlaTrHrValGlySerLeuAlaGlyInPro	220
Db	673	CCCCTGTGGGAACACAGGSCCGCTGGGGCCGCCCATCTGTGGCTCCCTGGCCAGCGCGG	732
OY	221	LeuGlnGluArGrAlaGlnAlaATrpGlyLuarGrLeuArGrAlaArGrMeGlnGluMet	239
Db	733	CTACAGGAGCGGGCCCGACGGCTGTGGGGCCAGCGGTG -CGCGCGGATGAGAGAGATG	788
RESULT 13			
B15S1475			
LOCUS			
DEFINITION	603194314P1 NIH_MGC_95 Homo sapiens CDNA clone IMAGE:5274003 5',		
ACCESSION	mRNA sequence.		
VERSION	B15S1475		
KEYWORDS	B15S1475.1 GI:15438787		
SOURCE	EST.		
ORGANISM	human.		
JOURNAL	Homo sapiens		
TITLE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 919)		
COMMENT	NIH-MGC http://mgs.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapds-remail.nih.gov Tissue Procurement: Miklos Palcovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki Toshiyuki and Piero Carninci (RIKEN) CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at: http://image.llnl.gov Plate: LLAM1691 row: 1 column: 04 High quality sequence stop: 812. Location/Qualifiers . . . 919 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5274003" /clone_lib="NIH_MGC_95" /cruise_type="hippocampus" /lab_host="DH10B"		
FEATURES			
source			

```

/note="Organ: brain; Vector: plusucscriptR (modified
pbpluscript Ks+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-gttttttttttttttttVN-3',
size-selected for average insert size 2.5 kb and
normalized to R0F 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carlnici, in preparation). Library
constructed by M. Brownstein (NIMH/NIHCR, National
Institutes of Health). Note: this is a NIH-MGC Library."
BASE COUNT      172 a      270 c      363 g      114 t
ORIGIN
Alignment Scores:
Pred. No.:      2.5e-109      Length:      919
Score:          1137.00      Matches:      246
Percent Similarity: 93.26%      Conservative: 3
Best Local Similarity: 92.13%      Mismatches:  17
Query Match:      71.92%      Indels:      4
DB:               13          Gaps:        0

```

DB	US-09-827-854-14 (1-317) x BI551475 (1-919)	
OY	1 MetLysValLeuThrPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
DB	73 ATGAAGATTCTGTGGCGTCCGTTGCTGTGCATTCCTGCAGAGATGCCAGGCCAAGGTG	132
OY	21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrGlnSer	40
DB	133 GAGCAAGCGGTGGAGACAGAGCGGAGGCCAGCGTCCGCCAGCAGACCGAGCTGGCAGAGC	192
OY	41 GlyGlnArgTrpGluLeuAlaLeuGlyIleArgPheTrpPheTrpPheLeuArgTrpValGlnThr	60
DB	139 GGGCAGCGCTGGAGACTGGCACTGGGTGGCTTTTGCGATTACCTGCGCTGGGTGCAGACA	255
OY	61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80
DB	253 CTGTCTGAGCAGGTCCAGAGAGAGCTGCTACGCTCCAGTCCACCCAGGAACTAGAGGGCG	312
OY	81 LeuMetAspGluThrMetLysGluLeuLeuLysAlaTrpLysSerGluLeuGluGluGlnLeu	100
DB	313 CTGAATGGACGAGAACCATAGAGAGTTGAAGCCCTTACAATGTGGAATCTGAGAACCACTG	372
OY	101 ThrProValAlaGluGluThrArgAlaIleArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
DB	373 ACCCGGCTGGCGGAGAAACCGCGGCGACGGCTGTCCAAAGAGCTGCAGCGCGCGACGGCC	433
OY	121 ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTrpArgGlyGluVal	140
DB	433 CGGCTGGCGCGGACATGAGAGACCTGTGGCGCCGCTGTGTCAGTACCGGCGGAGGGTG	492
OY	141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValAlaGluLeuAlaSerHisLeuArg	160
DB	493 CAGGCCATGCTCGGCGCAAGACCCGAGGAGGTGGGGGCGCCCTCGCTCCCACTCCG	553
OY	161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
DB	553 AAGCTGCGTAAAGCGCTCTCTCCGCGATCCCAATGACCTGCAGAAAGCGCTTGGCAGTAC	612
OY	181 GlnAlaGlyAlaArgGluGlyAlaGlnArgGlyLeuSerAlaIleArgGluArgLeuGly	200
DB	613 CAGGCGGGGGGCGGAGTGGCGGCGGTGCTGACACCGCATCCGAGACCGGCTGGGG	672
OY	201 ProLeuValGluGlnGlyArgValAlaArgAlaIleThrValGlySerLeuAlaGlyGlnPro	220
DB	673 CCCCTGGTGGAAACAGGCGCGGTGGGGGCCCACTGCTGGGCTCCCTGGCGAGGCTT	733
OY	221 LeuGlnGluArgGluAlaGlnAlaLeuTrpGlyGluArgLeuArgAlaArgMetGluGlu-MetGly	240
DB	733 A--CAGGAGCGGGGCGAGAGCTTGGGGCGAGGCGTGGCCGCGGATGGAGAAATGGG	790
OY	240 ySerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys	260
DB	791 CAGCCGGAGCCGGGACG--CTGGAGCGAAGTGAAGACACGTGGGCGAGAGG--TCCGCCAA	844

QY 260 sLeuGluglInlaGlIn 266
 DB 848 GCTGAGACACACAGAG 866

RESULT 14
 LOCUS BM728696 706 bp mRNA linear EST 01-MAR-2002
 DEFINITION UI-E-EJ0-aiu-1-08-0-UI.r1 UI-E-EJ0 Homo sapiens cDNA clone
 UI-E-EJ0-aiu-1-08-0-UI 5', mRNA sequence.
 ACCESSION BM728696
 KEYWORDS BM728696.1 GI:19050022
 EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 706)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.iowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse

FEATURES
 source
 Location/Qualifiers
 1..706
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="UI-E-EJ0-aiu-1-08-0-UI"
 /clone_1lb="UI-E-EJ0"
 /tissue_type="fetal eyes, lens, eye anterior segment,
 optic nerve, retina, Retina Foveal and Macular, RPE and
 Choroid"
 /dev_stage="fetal and adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UI-E-EJ0 is a subtracted cDNA library constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pT73-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tags for this library are: fetal eyes, AGATCAGAGA
 ; lens, CGATTACGGA; eye anterior segment, AATGCCGCAT;
 optic nerve, CCATTAGTG; retina, CCGCG; Retina Foveal and
 Macular, GTCC; RPE and Choroid, ACCGA. This library was
 created for the program, Gene Discovery in the Visual
 System, supported by National Eye Institute (NEI)."

BASE COUNT 123 a 218 c 283 g 82 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2.2e-109 Length: 706
 Score: 1136.00 Matches: 233
 Percent Similarity: 98.73% Conservative: 0
 Best Local Similarity: 98.73% Mismatches: 1

Query Match: 71.85% Indels: 2
 DB: 14 Gaps: 0
 US-09-827-854-14 (1-317) x BM728696 (1-706)

QY 15 GlyCysGlnAlaLysValGluGlnAlaVal-GluThrGluProGluProGluLeuArgG1 34
 DB 2 GGATGCCAGGCCAAAGGTGAGCAAGCGGTGAGAGAGAGCCGAGCCGAGCTGGCCA 61
 QY 34 ngIntnrglntprglnsergylglnargtrpctuleuAlaLeuGlyArgPheTrpAspTy 54
 DB 62 GCAGACCGAGTGGCAGAGCGGCCAGCGCTGGAGACTGGCAGCTGGCTTTGGATTA 121
 QY 54 rLeuArg-TrpValGlnThrLeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnV 74
 DB 122 CCGCCGCTGGGTCAGACACTGTGTGACAGAGTGCAGAGAGAGCTGTCACCTCCAG 181
 QY 74 alThrGlnGluLeuArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyLys 94
 DB 182 TCACCCAGAACATGAGGCGGTGATGACGAGACCATGAAAGGATGAAAGCCCTACACAT 241
 QY 94 erGluLeuGluGlnGlnLeuThrProValAlaGluGlnThrArgAlaArgLeuSerLysG 114
 DB 242 CGGAACCTGAGAACAACTGACCCCGGTGGAGAGAGACCGCGGCTGTCCAAAG 301
 QY 114 ludeGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuV 134
 DB 302 AGCTGAGCGGCGGCGAGCCCGGCTGGCGGACATGAGAGACGTCGCGCCGCTCG 361
 QY 134 alGlnTrpArgGlyGluValGlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgVal 154
 DB 362 TGAGTACCGCGCGGAGGTGAGGCGCATGTGCGCCAGACACCGAGAGCTCGGGTCC 421
 QY 154 rgluAlaSerHisLeuAlaGlyLysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuG 174
 DB 422 GCGTGGCTCCACACCTGGCCAGAGCTCGTAAGGGCTCTCCGCGATGGCATGACTCC 481
 QY 174 lnlYsArgLeuAlaValTyrGlnAlaGlyAlaArgGluGluValArgGlyLeuSerA 194
 DB 482 AGAAGCGCTGGCAGTGCACAGGCGGCGGCCGCGCAGGCGCGGCTCTCAGG 541
 QY 194 lalleArgGluArgLeuGluProLeuValGluGlnGlnValArgAlaAlaThrValG 214
 DB 542 CCATCCGCGAGGCGCTGGGGCCCTGTGTGAAACAGGGCCGCGGCGGCGGCTGTGG 601
 QY 214 lYserLeuAlaGlnProLeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgA 234
 DB 602 GCGTCTGGCGCGGCCGCGCTACAGAGCGGGCCAGGCGCTGGGCGAGCGGCTGGCG 661
 QY 234 laArgMetGluMetGlySerArgThrArgAspArgLeuAsp 248
 DB 662 CCGGATGAGAGATGGGCGACCGGACCGCGGCGGCGGCTGGAGC 705

RESULT 15
 LOCUS AK010261 1110 bp mRNA linear HTC 19-JAN-2002
 DEFINITION Mus musculus ES cells cDNA, RIKEN full-length enriched library,
 clone:2400003C12;apolipoprotein E, full insert sequence.
 ACCESSION AK010261
 KEYWORDS AK010261.1 GI:12845571
 SOURCE HTC, CAP trapper.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Carninci,P. and Hayashizaki,Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE AUTHORS	TITLE	JOURNAL MEDLINE PUBMED	REFERENCE AUTHORS	TITLE	JOURNAL MEDLINE PUBMED	
2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	Genome Res. 10 (10), 1617-1630 (2000)	3	Shibata, K., Itoh, M., Alizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujisake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipillary sequencer	Genome Res. 10 (11), 1757-1771 (2000)	
4	Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Haru, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Alizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Konno, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kodato, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Knehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staudt, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hoffmann, M., Hume, D. A., Kamlay, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyok, O. K., Wang, K. H., Weitz, C., Wittaker, C., Wilming, L., Wyszewski, B., Yoda, K., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S. and Hayashizaki, Y.	Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)	5	(bases 1 to 110)	Adachi, J., Alizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Haru, A., Hayatsu, N., Hill, D., Hironaka, K., Hiraoka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koyu, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sugabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyok, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.	Direct Submission Submitted (10-JUL-2000) Yoshitake Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome-gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	Please visit our web site (http://genome-gsc.riken.go.jp/) for further details. GSC library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken.					

```

Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAAGACGATCCAGAGCTCTTTTGTTCCTTTTTCNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. Second
strand cDNA was prepared with the primer adapter of sequence [5'
GGAGGAGCATCTCAGCAATTAAATAAATTAATGCCCCCCCCC 3']. cDNA was cleaved
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.

FEATURES             location/Qualifiers
source                1..1110
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:2400003C12"
                     /db_xref="MGI:1904659"
                     /db_xref="taxon:10090"
                     /clone="2400003C12"
                     /cell_type="ES cells"
gene                 1..1110
                     /clone_id="RIKEN full-length enriched mouse cDNA library"
misc_feature          1..1110
                     /gene="Apoe"
                     /note="apolipoprotein E
                     data source:MGI, source key:MGI:88057, evidence:ISS"
BASE COUNT           260 a      309 c      362 g      179 t
ORIGIN
Alignment Scores:
Pred. NO.:            6.76e-109              Length:        1110
Score:               1134.00                  Matches:         223
Percent Similarity:   82.39%                   Conservative:    39
Best local Similarity: 70.13%                   Mismatches:     46
Query Match:          71.73%                   Indels:         10
DB:                   11                      Gaps:           2

US-09-827-854-14 (1-317) x AK010261 (1-1110)
OY       1 MettysValleuuTPPAlalaLeuLeuValThrPhereuAlaglyCySGlnAlaLysVal 20
          ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       67 ATGAAGCGCTCTGTGGGCCCTGCTGTTGTGCATATGCTGACAGAGTGCTGAGCC----- 120
OY       21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       121 -----GAGGAGAGCGCGGAGGAGGAGGACAGATCAAGTCAGCTGAGCAAGC 162
OY       41 GlyGlnArGrPrGluLeuAlaLeuGlyArPherThrPasprYLLeuArGrPrValGlnThr 60
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       163 AACCAACCCTGGGAGCAGAGCCCTGMAACCGCTTGTGGATTCCTGCGCTGGTGAGACG 222
OY       61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgLa 80
          ||||| : | | | | | | | | | | | | | | | | | | | | | | | |
Db       223 CTGTGTGCACAGGTCCAGGAAGAGCTGCAGAGCTCCCAAGTCACACAAGACTGACGGCA 282
OY       81 LeuMetAspGluThrMetLysGluLeuLysAlaTryLysSerGluLeuGluGluGlnLeu 100
          ||| : | | | | | | | | | | | | | | | | | | | | | | | |
Db       283 CTGATGCGAGCACACATATGACGGAAGTAAAGGCTTAAACAAGAGAGCTGAGAACACGTG 342
OY       101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuAlaAlaGlnAla 120
          ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db       343 GGTCCTCAGTGGCGGAGGAGACACGGGCCAGGCTGGGCAAGAGGTGCACAGCGCACAGGCC 402
OY       121 ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTryArgGlyGlnVal 140
          ||||| : | | | | | | | | | | | | | | | | | | | | | | | |
Db       403 CGACTCGGAGCCGACATGAGATCTACGCCAACCGACTCGGGCAGTACCGCAACGAGGTG 462
OY       141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
          ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db       463 CACACCAATGCGTGGGCGCAGAGCACAGAGAGATACGGGGCGGCTTCCACACACTGGCC 522
OY       161 LysLeuArGlySarGleuLeuArGAspAlaAspAspLeuGlnLysArgLeuAlaValTrp 180
          ||| : | | | | | | | | | | | | | | | | | | | | | | | |

```



```
Db 523 AACATGCCCAAGCCCTTGATGCGGGATGCCAGAGATCTGCAGAACGCCCTAGCTGTGTAC 582
QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 583 AAGCAGGGGAGCGCTAGGGGCGCGAGCGCGGTGTGAGTGCCATCCGTGAGCGCCTGGGG 642
QY 201 ProLeuValGluGlnGlyArgValArgAlaIleThrValGlySerLeuAlaGlyGlnPro 220
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 643 CCTCTGGTGGAGCAAGGTGCGCCAGCGCATGCGCAACCTAGGCGCTAGGCGCGCCAGACCT 702
QY 221 LeuGlnGluArgAlaGlnAlaIleTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 703 ATGCGCATGCGCGCCAGCGCTTTGGTGACCGCATCCGAGGGCGGCTGGAGCAAGTGGGC 762
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 763 AACCAAGGCCGCGTAGCCGCTAGAGAGAGTGGGTGACACATGAGCAGAGTGGCTCCAG 822
QY 261 LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 823 ATGAGAGACAGACAGCCCAATATACGCTGACGGGAGATCTTCCAGGCCCGCCTCAAG 882
QY 281 SerTrpPheGluProLeuValGluAspMetGluArgIleTrpAlaGlyLeuValGluLys 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 883 GGCTGTTCCAGCCCAATAGTGAGACATGCATCGCCAGTGGCCAAACTGATGGAAGAAG 942
QY 301 ValGlnAlaAlaValGlyThrSer-----AlaAlaProValProSerAspAsn 316
   ::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 943 ATACAGGCGCTGTGCTACCAACCCCATCATCATCCCAAGTGGCCGAGAGAAAT 996
```

Search completed: March 14, 2003, 20:13:47
Job time : 1296.69 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:08:57 ; Search time 2307.23 Seconds
(without alignments)
3998.551 million cell updates/sec

Title: US-09-827-854-15
Perfect score: 1585
Sequence: 1 MKVLMALLVTLGACQAKV.....VEKVQAVGTSAPVPSDNH 317

Scoring table: BLOSUM62
Xgapop 10.0, Ygapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 segs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODL=frame_plus_p2n model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09827854/rnact_11032003_101610_27486/app_query.fasta_1.3576
-DB=genembl -QFMT=fastlap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=plc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09827854.cgn 1.1 13008 @rnact_11032003_101610_27486 -NCPU=6 -ICPU=3
-MARK_TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Genembl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_pl:*
25: em_pi:*
26: em_to:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1585	100.0	1110	6	E00359
2	1585	100.0	1110	6	E00823
3	1585	100.0	1147	6	AX302545
4	1585	100.0	1156	6	BD004278
5	1585	100.0	1156	9	HUMABOE3
6	1585	100.0	1186	9	BC003557
7	1578	99.6	1110	6	E08423
8	1577	99.5	1156	6	BD004277
9	1576	99.4	1157	6	AX333278
10	1576	99.4	1157	6	AX409597
11	1576	99.4	1157	6	115975
12	1576	99.4	1157	6	HUMABOE
13	1573	99.2	1156	6	BD004279
14	1451	91.5	1178	9	MFAPOE
15	1396.5	88.1	5491	9	AF261279
16	1396.5	88.1	41907	6	AX358722
17	1396.5	88.1	41907	9	AF050154
18	1396.5	88.1	107567	9	AC011481
19	1379	87.0	5515	9	HOMABOE4
20	1350	85.2	5413	9	AF261280
21	1267	79.9	4762	9	BABAPOE
22	1266	79.9	208239	2	AC021988
23	1258	79.4	1138	4	AF303830
24	1170	73.8	1060	4	RABAPOLP
25	1163.5	73.4	965	6	AX384545
26	1163.5	73.4	1108	4	BTAPOLPE
27	1163.5	73.4	5617	6	AX384541
28	1163.5	73.4	6026	6	AX384539
29	1151	72.6	718	9	AF20049783
30	1149	72.5	1045	10	MUSAPOE
31	1149	72.5	1104	10	BC028816
32	1147	72.4	718	9	AF20050053
33	1147	72.4	718	9	AF20050353
34	1145.5	72.3	1154	4	BRAPOEHR
35	1143	72.1	718	9	AF20050653
36	1143	72.1	1122	4	SSAPOE
37	1138	71.8	1126	6	AR164342
38	1138	71.8	1126	6	AR205885
39	1094	69.0	951	10	S76779
40	1072	67.6	959	10	MUSAPOE
41	1056.5	66.7	1069	10	RATPOE
42	1029	64.9	228698	2	AC127479
43	1029	64.9	237653	2	AC127479
44	1003	63.3	4856	10	MUSAPOE
45	1002	63.2	4267	6	AR164387

RESULT 1

ALIGNMENTS

LOCUS	E00359	1110 bp	RNA	linear	PAT 29-SEP-1997
DEFINITION	cDNA coding human apolipoprotein E3.				
ACCESSION	E00359				
VERSION	E00359.1	GI:2168646			
KEYWORDS	JP 1985118189-A/1.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens.				
REFERENCE	Mikayvota, Matsuza; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1110)				
AUTHORS	Teranishi, Y., Takamatsu, N., Matsui, Y., Kimura, M. and Ikeda, Y.				
TITLE	DNA FRAGMENT.				
JOURNAL	Patent: JP 1985118189-A 1 25-JUN-1985;				
COMMENT	MITSUBISHI CHEM IND LTD				
	OS	human			

FEATURES	source	Location/Qualifiers	product='apolipoprotein E3',	mat_peptide	FT
BASE COUNT	198 a	353 c	416 g	143 t	
ORIGIN					

Alignment Scores:	
Pred. No.:	9.31e-31
Score:	1585.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	6
Length:	1110
Matches:	317
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

[illegible]

OY	101	ThrpVAlaIAGLUGLnThrArgLaArgrLeuSerLySGluLeuGlnAlaIAGlnAla	120
Db	315	ACCCCGGTGGCGAGAGAACGCCGGCACCCTGTCTCCAAAGACTGCAGCGGCCAAGGCC	374
OY	121	ArgLeuGIyAlaAspMetGLuaSPvAlCySGIyArgLeuValGlnTyrArgSlyGlVal	140
Db	375	CGGCTGGGGCGGCAATGAGAGACGTGTCCGGCCGCTGTGCAGTAACCGCGGCGAGTG	434
OY	141	GlnAlaMetLeuGIyGlnSerThrTrglutLuLeuArqValArgLeuAlaserHisLeuArg	160
Db	435	CAGGCCATGCTCGCGCACAGCACACGAGAGAGCTGGCGGCTGCCCTCCACCTCGGCC	494
OY	161	LysLeuArqLyArGrLeuLeuArqSPvAlaSPaSPleuGlnLyArqLeuAlaValTyr	180
Db	495	AAGTCGCGTAAcGGGCTCTCCGCGCATGCGGTGAACCTGCACAACCGCTTGGCAGGTAC	554
OY	181	GlnAlaGIyAlaArGIuGIyAlaGIuArgGIyLeuSerAlaIlleaRgSIuArgLeuGIy	200
Db	555	CAGGCCGGGGCCCCGGAGGGCCCCCAAGCCGGGCTTCAGCGCATTCGGCAGCGCCTGGGG	614
OY	201	ProLeuValGIuGIuGIyArgValArqAlaArhArValGIySerLeuAlaGIyGlnPro	220
Db	615	CCCCTGTGTGAACAGGGCGCGCTGGCGGGCGCCACTGTGGCTCCCTGGCCGGCAGCGG	674
OY	221	LeuGIuGIuArqAlaGIuAlaArTrpLyGIuArgLeuArqAlaArqMeGIuGluMetGIy	240
Db	675	CTACAGGAGCGGGGCCACAGGCTGTGGGGCAGCGCGTGGCGCGGATGAGAGATGGGC	734
OY	241	SerArqTrrArqASpArgLeuaspGIuValLySGluGIuAlaGIuAlaGIuValArqAlaLyS	260
Db	735	AGCCGGACCCCGACGCCGCTGTGACCAAGGTGAAGAAGCAGTGGCGGAGTGGCCCCAAG	794
OY	261	LeuGIuGIuGIuAlaGIuGlnIleArgLeuGlnAlaGIuAlaIapheGlnAlaArgLeuLyS	280
Db	795	CTGGAGGAGACAGGCCACAGATACGCCCTGCAGGCGGAGCCTTCAGAGCCCGGCTCAAG	854
OY	281	SerTrpPheGIuProLeuValGIuaspMetGlnArqGIuINTPrAlaGIyLeuValGIuLyS	300
Db	855	AGCTGGTTCTGAGCCCTGTGTGGAAACATGTGACGCCACAGTGGCGGGCTGTGTGAGAG	914
OY	301	ValGlnAlaAlaValaIGlYThrSerAlaIAProValProSerASPasnHis	317
Db	915	GTGCAGGCTGCCGTGTGGCACCAAGCCGCCCTGTGTGCCCAGCGACAAATCAC	965
RESULT 2			
E00823		1110 bp DNA linear PAT 29-SEP-1997	
LOCUS			
DEFINITION	E00823	DNA sequence coding for human apolipoprotein E and its signal peptide.	
ACCESSION	E00823		
VERSION	E00823.1	GI:2169084	
KEYWORDS	JP 1986096997-A/1.		
SOURCE	unidentified.		
ORGANISM	unidentified.		
REFERENCE	unclassified.		
AUTHORS	1 (bases 1 to 1110)		
TITLE	Teranishi,Y., Matsui,Y., Ikeda,Y. and Kimura,M.		
JOURNAL	PRODUCTION OF HUMAN APOLIPOPROTEIN E-LIKE PROTEIN Patent: JP 1986096997-A 1 15-MAY-1986;		
	MITSUBISHI CHEM IND LTD		
COMMENT	OS Human (Homo sapiens) PN JP 1986096997-A/1 PD 15-MAY-1986 PF 16-OCT-1984 JP 1984216987 PI TERANISHI YUTAKA, MATSUJI YASUSHI, IKEDA YASUOKO, KIMURA MASAKO PC C12P21/00,A61K35/74,A61K37/04,C12N15/00,(C12P21/00,C12N1:19), PC (C12N15/00), PC C12N1:19); CC strandedness: Double; CC topology: Linear; CC hypothetical: No; CC anti-sense: No; CC *source: tissue_type=Liver;		

FH	Key	Location/Qualifiers
FH	3'UTR	1..14
FT	sig_peptide	15..68
FT	/product='human apolipoprotein E signal'	FT
FT	peptide	69..968
FT	CDS	/product='human apolipoprotein E' FT 3'UTR
FEATURES	source	location/Qualifiers 1..1110 /organism="unidentified" /db_xref="taxon:32644"
BASE COUNT	198 a 353 c 416 g 143 t	
ORIGIN		
Alignment Scores:		
Pred. No.:	9.31e-81	Length: 1110
Score:	1585.00	Matches: 317
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	Gaps: 0	
US-09-827-854-15 (1-317) x E00823 (1-1110)		
OY	1 MetysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	15 ATGAAGTTCTGTGGCGTCTGTCACATTCCTGGCAGAGATGCCAGGCCAAGTG	74
OY	21 GUGGlnAlaValAlaGluThrGluProGluLeuArgGlnGlnThrGluThrGlnSer	40
Db	75 GAGCAAGCGGTGGAGACAGACCAGCCAGCTGGCCAGACAGACGAGTGGCAGAC	134
OY	41 GlyAlnArGrTPGluLeuAlaLeuGlyAraPhetPrAspyrLeuArGrTPValGlnThr	60
Db	135 GGCCAGCGCTGGAGACTGGCTGGCTTTGGATTAACCTGGCTGGTGCGAGACA	194
OY	61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80
Db	195 CTGCTGAGCAGGTGCAGAGAGACTGCTCACAGTCCAGAGTCCAGAGAACTGAGGGCG	254
OY	81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrltySerGluLeuGluGlnGlnLeu	100
Db	255 CTGATGGACGAGACCATGAAGAAGTGAAGCTTACAAATCGGAGACTGGAGAACAACTG	314
OY	101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	315 ACCCGGTGGCGGAGAGACCGCGGCGACGGCTGTCCAGAGAGCTGGAGGGGGCGAGGCC	374
OY	121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal	140
Db	375 CGCGTGGCGGCGGAGCATGAGAGAGTGTGGCGCGCTGGTGCATGACCGCGCGAGG	434
OY	141 GlnAlaMetLeuGlnInsThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	435 CAGCGCATGCTCGCGCCAGAGACCGAGAGACTCGGGTGGCTCCCTCCACCTCGCGC	494
OY	161 LysLeuArgLysAspLeuLeuArgAspAlaAspAspLeuLysArgLeuAlaValTyr	180
Db	495 AAGCTGCTAAGCGGCTCTCCCGGATGCGGATGACCTGCAAGAGCGCTGGAGGTATC	554
OY	181 GlnAlaGlyAlaAspGluGlnGlyAlaGluArgGlyLeuSerSerAlaIleArgGluArgLeuGly	200
Db	555 CAGCGCGGGCGCGCGAGGGCGCGAGCGGCTCAGCGCCATCCGCGAGCGCTGGGG	614
OY	201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValCylserLeuAlaGlyGlnPro	220
Db	615 CCCCTGGTGGTAAAGCGCGCTCGGGCGCGCACTGGGCTCCCTGGCGCGCGCGCG	674
OY	221 LeuGlnGluArgGlnAlaAlaTrpGlyGluArgLeuArgAlaArgMetGluGlnMetGly	240
Db	675 CTCACGGCG	734

OY	241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys	260
Db	735 AGCCGAGACCCCGACCGCTGGAGAGAGTAGAGAGAGTGGAGGATGGCGCGCGCAAG	794
OY	261 LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaGlnAlaGlnAlaGln	280
Db	795 CTGAGAGAGAGCGCGCGAGATGAGCTTCAGAGCGAGCGCTTCAGCGCGCGCGCAAG	854
OY	281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys	300
Db	855 AGCTGCTTCAGAGCGCGCTGGTGAAGATGACAGCGCGAGGGCGGCTGGTGGAGAAG	914
OY	301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis	317
Db	915 GTGACAGCTGCGGTGGCGCACAGCGCGCGCGCTGTCAGCAGCAGCAATCAC	965
RESULT 3		
AX302545	AX302545	1147 bp DNA linear PAT 30-NOV-2001
LOCUS	Sequence 63 from Patent WO0175177.	
DEFINITION	AX302545	
ACCESSION	AX302545	
VERSION	AX302545.1 GI:17383082	
KEYWORDS		
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 Morin,P.J., Sherman-Baust,C.A., Pizer,E.S. and Hough,C.D. TITLE Tumor markers in ovarian cancer Patent: WO 01/5177-A 63 11-OCT-2001; JOURNAL THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US) FEATURES location/Qualifiers source 1..1147 /organism="Homo sapiens" /db_xref="taxon:9606"	
BASE COUNT	210 a 365 c 425 g 147 t	
ORIGIN		
Alignment Scores:		
Pred. No.:	9.64e-81	Length: 1147
Score:	1585.00	Matches: 317
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	Gaps: 0	
US-09-827-854-15 (1-317) x AX302545 (1-1147)		
OY	1 MetysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	46 ATGAAGTTCTGTGGCGTCTGTCACATTCCTGGCAGAGATGCCAGGCCAAGTG	105
OY	21 GUGGlnAlaValAlaGluThrGluProGluLeuArgGlnGlnThrGluThrGlnSer	40
Db	106 GAGCAAGCGGTGGAGACAGAGCGAGACCGAGACTGGCCAGACAGACCGAGTGGCAGAC	165
OY	41 GlyAlnArGrTPGluLeuAlaLeuGlyAraPhetPrAspyrLeuArGrTPValGlnThr	60
Db	166 GCCAGCGCTGGAGAACGTGGAGCTGGGTGGTGGATTAACCTGGCTGGGGCGAGACA	225
OY	61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80
Db	226 CTGCTGAGCAGAGTGCAGAGAGACTGCTACGCTCCAGGTACCCAGGAAGTGAAGGCG	285
OY	81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrltySerGluLeuGluGlnGlnLeu	100
Db	286 CTGATGGACGAGACCATGAAGAAGTGAAGCTTACAAATCGGAATCGAGAACAACTG	345
OY	101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	346 ACCCGGTGGCGGAGAGCGCGCGCGCTGTCCAGAGAGCTGGAGCGCGCGCGCGCG	405

QY 121 ArgLeuGIyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 DB 406 CGGCTGGGGCGGACATGAGAGACGTGTGCGCGCTGTGTGAGTACCGCGGAGAGT 465
 QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 466 CAGGCGCATCTGCGGCGAGACACCGAGAGCTCGGGTCCGCTCCCTCCACCTGCGG 525
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 DB 526 AAGCTCGTAAGGAGCTCTCCGCGATGCCGATGACCTCAGAAAGCCCTGGCGAGTGTAC 585
 QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
 DB 586 CAGCGCGGGGGCGCGGAGGGCGCGGAGCGCGGCTCAGCGGCACTCCGCGAGCGCTGGG 645
 QY 201 ProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
 DB 646 CCCCTGGTGGAAACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 705
 QY 221 LeuGlnGlnArgAlaGlnAlaTrrpGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
 DB 706 CTACAGAGACGGGCGGACAGGCGCTGGGCGAGCGGCTGCGCGCGGAGTGAAGATGGGC 765
 QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValAlaArgAlaLys 260
 DB 766 ACCCGGACCGCGGACCGCTGGAGCGAGTGAAGAGACAGTGGCGGAGGTGGCGCGCG 825
 QY 261 LeuGlnGlnGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
 DB 826 CTGAGAGACAGGCGCGGACGATACCGCTGCGAGCGCGCGCTTCCAGCGCGCGCGCAAG 885
 QY 281 SerTrrpPheGlnProLeuValGlnAspMetGlnArgGlnTrrpAlaGlyLeuValGlnLys 300
 DB 886 AGCTGTTGAGACCCCTGTGTGAGACATGACAGCGGCAATGGGCGCGGCTGTGGAGAG 945
 QY 301 ValGlnAlaAlaValGlnGlnSerAlaAlaProValProSerAspAsnHis 317
 DB 946 GTGCAGAGTGCCTGTGGCGACAGCGCGCGCTGTGCCAGGACAAATCAC 996
 RESULT 4
 LOCUS BD004278 1156 bp DNA linear PAT 31-JAN-2002
 DEFINITION Apo E humanized mammal.
 ACCESSION BD004278
 VERSION BD004278.1 GI:18632239
 KEYWORDS JP 2001017028-A/2.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.
 TITLE Apo E humanized mammal
 JOURNAL Patent: JP 2001017028-A 2 23-JAN-2001;
 MITSUBISHI CHEMICAL CORP
 OS Homo sapiens (human)
 PN JP 2001017028-A/2
 PD 23-JAN-2001
 PF 28-APR-2000 JP 2000128919
 PR
 PI SHINOBU FUJITA,HIROKI HAMANAKA,YUKO FUKUI,MINESUKE YOKOYAMA PC
 A01K67/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10, PC
 C12N15/09,CO7K14/775,
 PC (C12N5/10,C12R1:91),C12N5/00,(C12N15/00,(C12N5/00,C12R1:91) CC
 FEATURES
 source FT key CDS Location/Qualifiers
 1..1156 Location/Qualifiers
 /db_xref="taxon:9606"

BASE COUNT 208 a 368 c 432 g 148 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 9,71e-81 Length: 1156
 Score: 1585.00 Matches: 317
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-827-854-15 (1-317) x BD004278 (1-1156)
 QY 1 MetLysValLeuTrrpAlaAlaLeuLeuValTrrpPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 61 ATGAAGATTCTGTGGGCTGCTGTTGCTGTGACATTCTCTGGAGAGATGCCAGGCAAGTGC 120
 QY 21 GlnGlnAlaValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 40
 DB 121 GAGCAAGCGGTGGAGACAGAGCGGAGCGGAGCGGAGCTGCGGAGAGCGAGTGGCAGACC 180
 QY 41 GlyGlnArgTrrpGlnLeuValLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 60
 DB 181 GGCGACCGCTGGGAGACTGACCTGGGCTCTTTGGGATTTACTGCGCTGGTGCAGACA 240
 QY 61 LeuSerGlnGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
 DB 241 CTGTCTGACAGAGTGCAGAGAGAGTGTCTCAGCTCCAGGTCACCCAGGACTGAGAGGCG 300
 QY 81 LeuMetAspGlnThrMetLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
 DB 301 CTGATGAGAGAGACCAATGAAGAGGTTGAAGGCTTAAATTCGAAATCGAGAGAACAACTG 360
 QY 101 TrrpProValAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120
 DB 361 ACCCGGCTGGCGAGAGAGCGGCGGACGCTGTCTCCAGGACTGCGAGCGCGCGGAGGCC 420
 QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 DB 421 CGGCTGGGGCGGACATGAGAGACGTGTGCGCGCGCTGTGTGAGTACCGCGGAGAGT 480
 QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
 DB 481 CAGGCGCATCTCTGCGGCGAGACCGGAGGCTGCGGCTGCGCTGCCCTCCACCTGCGGC 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 DB 541 AAGCTGCGTAAGCGGCTCTCCGCGATGCGGATGACCTCAGAGCGCTGCGCAGTGTAC 600
 QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
 DB 601 CAGGCGGGGGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 660
 QY 201 ProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
 DB 661 CCCCTGGTGGAAACAGGCGCGGCTGGGCGGCGGCGGCTGTGGCTGCTGCGCGCGCGG 720
 QY 221 LeuGlnGlnArgAlaGlnAlaTrrpGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 240
 DB 721 CTACAGAGACGGGCGGAGGCGCTGGGCGGAGCGGCTGCGCGCGGATGAGAGATGGGCG 780
 QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValAlaArgAlaLys 260
 DB 781 ACCCGGACCGGAGCGGCTGTGAGCGAGTGAAGAGACAGTGGCGAGGTGGCGCGCAAG 840
 QY 261 LeuGlnGlnGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
 DB 841 CTGAGAGAGACGCGGACGAGATACCGCTGCGAGCGGAGCGCTTCCAGCGCGCGCTCAAG 900
 QY 281 SerTrrpPheGlnProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 300
 DB 901 AGCTGTTGAGACCCCTGTGTGAGACATGACAGCGGCAATGGGCGGCGGCTGTGGAGAGAG 960

Db	Accession	Definition	Version	Keywords	Source	Organism	Reference	Authors	Title	Journal	Medline	PubMed	Comment
301	ValginalaAlaValaIGlyThrSerIleAlaProValProSerAspAsnHis 317												
961	GTCCAGGCTGCCGTGGCACCAGCGCCGCCCTGTGTGCCAGCAGACATCAC 1011												
RESULT 5	HUMAPOE3	1156 bp	mRNA	linear	PRI 24-NOV-2000								
LOCUS	HUMAPOE3	Homo sapiens preapolipoprotein E (APOE) mRNA, complete cds.											
DEFINITION	Homo sapiens preapolipoprotein E	apolipoprotein E; lipoprotein; polymorphism; very low density lipoprotein.											
ACCESSION	K00396	G1:178850											
VERSION	K00396.1	G1:178850											
KEYWORDS													
SOURCE	Homo sapiens												
ORGANISM	Homo sapiens												
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.												
AUTHORS	1 (bases 355 to 1156)												
TITLE	Breslow,J.L., McPherson,J., Nussbaum,A.L., Williams,H.W.,												
JOURNAL	Identification and DNA sequence of a human apolipoprotein E cDNA												
MEDLINE	J. Biol. Chem. 257 (24), 14639-14641 (1982)												
PUBMED	83082756												
AUTHORS	6897404												
REFERENCE	2 (bases 250 to 777)												
AUTHORS	Wallis,S.C., Rogne,S., Gill,L., Markham,A., Edge,M., Woods,D.,												
TITLE	Williamson,R., and Humphries,S.												
JOURNAL	The isolation of cDNA clones for human apolipoprotein E and												
MEDLINE	detection of apoE RNA in hepatic and extra-hepatic tissues												
PUBMED	EMBO J. 2 (12), 2369-2373 (1983)												
AUTHORS	84131952												
REFERENCE	3 (bases 1 to 1156)												
AUTHORS	Zannis,V.I., McPherson,J., Goldberger,G., Karathanasis,S.K. and												
TITLE	Breslow,J.L.												
JOURNAL	Synthesis, intracellular processing, and signal peptide of human												
MEDLINE	apolipoprotein E												
PUBMED	J. Biol. Chem. 259 (9), 5495-5499 (1984)												
AUTHORS	84185684												
REFERENCE	4 (bases 88 to 1156)												
AUTHORS	McLean,J.W., Eshourbagy,N.A., Chang,D.J., Mahley,R.W. and												
TITLE	Thayer,J.M.												
JOURNAL	Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing												
MEDLINE	of a new variant												
PUBMED	J. Biol. Chem. 259 (10), 6498-6504 (1984)												
AUTHORS	84212473												
REFERENCE	5 (bases 577 to 624)												
AUTHORS	Gill,L.L., Peoples,O.P., Pearson,D.H., Robertson,F.W.,												
TITLE	Humphries,S.E., Cumming,A.M. and Hardman,N.												
JOURNAL	Isolation and characterisation of a variant allele of the gene for												

epsilon-3 and epsilon-2. This sequence appears to be of the
 epsilon-3 allele. [1] argues that the apo E polymorphism involves
 mutations in the structural coding region; for example the
 epsilon-2 phenotype which is characterized by hyperlipoproteinemia
 is thought to result from a c to t change (arg to cys) at base 586
 below [31],[5]. The sequence shown is 57% homologous with human apo
 A-1 and 81% homologous with rat apo E. For the epsilon-4 sequence,
 see the separate entry
 [J. Biol. Chem. 258, 11422-11422 (1983)] is too extensively revised
 relative to [1] to record the
 revisions in the FEATURES table. The sequence below is that of [J.
 Biol. Chem. 258, 11422-11422 (1983)]
 and [3]:
 Apo E is located on chromosome 19 --Jackson, Bruns and Breslow ,
 PNAS USA 81, 2945-2949 (1984) -- and is believed to be linked to
 the Apo C-II gene (see separate entry).
 The two mutations causing type II hyperlipoproteinemia
 (apolipoprotein E phenotype E3/3) produces substitutions of Arg for
 Cys at amino acid 112 and Cys for Arg at amino acid position 142.
 Draft entry and printed copy of sequence for [1] kindly provided by
 L.L. Gill, University of Aberdeen.
 Complete source information:
 Human liver [1], [J. Biol. Chem. 258, 11422-11422 (1983)], [2], [3],
 [4] and blood [5], cDNA to mRNA.
 Location/Qualifiers
 1..1156
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="19q13.2"
 /tissue_type="liver and blood"
 1..1156
 /gene="APOE"
 61..1014
 /gene="APOE"
 /codon_start=1
 /product="preapoliipoprotein E"
 /protein_id="AAB59346.1"
 /db_xref="GI:178851"
 /db_xref="GDB:G00-119-691"
 /translation="MKVIMAAALVTFLAGCOAKVEAOVETPEPELHQETEMOSGORM
 EIALGRWDYLRWQTLSEVOQEFLLSSQYQOELRALMDETMKLKYKSLEROLNP
 VAEETRRLRSKELOAAQARIGADMEDYCGRLVQYRGEVQALGSGTEFLRRLASHLR
 KLRRLRLADDDLLQRLAVYQAGAREGEAERGLSIRRELPVLEQGVRAVAVGSLAQ
 QPRLRQAQAGERLRARMEMGSRTRRLDEVRKQVAVERAKLEEQAAQQLRQLAEAFQ
 ARLSAFPEPLVEDMQRQWAGIVEKYQAAVGTSAPVPSDNH"
 61..114
 /gene="APOE"
 /note="apoliipoprotein E"
 115..1011
 /gene="APOE"
 /product="apoliipoprotein E mature peptide"
 267
 /gene="APOE"
 /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)]. [3],
 [4]"
 /citation=[2]
 /replace="c"
 279
 /gene="APOE"
 /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)]. [3]"
 /citation=[4]
 /replace="a"
 288
 /gene="APOE"
 /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)]. [3]"
 /citation=[4]
 /replace="a"
 369
 /gene="APOE"
 /note="in variant"
 /citation=[4]
 /replace="a"
 409


```

/ gene="APOE"
/ note="in variant"
/ citation=[4]
/ replace="a"
448
variation
/ gene="APOE"
/ note="in type III hyperlipoproteinemia"
/ replace="c"
538
variation
/ gene="APOE"
/ note="in type III hyperlipoproteinemia"
/ replace="t"
568
variation
/ gene="APOE"
/ note="in variant"
/ citation=[4]
/ replace="c"
586
variation
/ gene="APOE"
/ note="in epsilon-2 allele"
/ citation=[3]
/ citation=[5]
/ replace="t"
702
variation
/ gene="APOE"
/ note="c in [J. Biol. Chem. 258, 11422-11422 (1983)], [3], [4]"
/ citation=[2]
/ replace="g"
783
variation
/ gene="APOE"
/ note="in variant"
/ citation=[4]
/ replace="t"
858
variation
/ gene="APOE"
/ note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3] and [4] variant; a in [4] normal,"
/ citation=[4]
/ replace="a"
BASE COUNT      208 a      368 c      432 g      148 t
ORIGIN
Alignment Scores:
Pred. No.:      9,71e-81      Length:      1156
Score:          1585.00      Matches:      317
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:              9      Gaps:      0
US-09-827-854-15 (1-317) x HUMAPOE3 (1-1156)
QY      1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
      |||||||
Db      61 ATGAAGGTTCTGGGGCTGGTGGCTGCATCTCTGGCAGAGTCCGAGCCAGAGTG 120
QY      21 GlnGlnAlaValGlnTrpGlnProGlnProGlnLeuArgGlnInPheGlnTrpGlnSer 40
      |||||||
Db      121 GACCAAGCGGTGGAGACAGACCGGACCGGACCTGGCCAGACAGACCGAGTGGCAGAC 180
QY      41 GlnGlnArgTrpGlnLeuAlaLeuGlnLysArgPheTrpAspTyrLeuArgTrpValGlnThr 60
      |||||||
Db      181 GGCACAGCGCTGGGAACTGGGACTGGCTGCTTTGGGATTACCTGGCGCTGGTGAGACA 240
QY      61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
      |||||||
Db      241 CTGCTGAGCAGGTGGAGGAGGAGCTGCATCCAGGTACCCAGGAACTGAGAGGCG 300
QY      81 LeuMetAspGlnTrpMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnGlnLeu 100
      |||||||
Db      301 CTGATGAGCAGACCAATGAAGAGTTGAAGGCTTACAATCGAAGTGGAGGACAAC 360

```

```

QY      101 ThrProValAlaGlnGlnTrpArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
      |||||||
Db      361 ACCCGGTGGCGGAGAGGAGCGGGCCAGCGCTGTCCAAAGAGCTGACAGCGGGCGAGGCC 420
QY      121 ArgLeuGlnValAspMetGlnAspValCysGlyArgLeuValGlnTrpArgGlyVal 140
      |||||||
Db      421 CGGCTGGCGCGGACATGAGAGAGTGTGGCCGCTGTGTGACATACCAGGCGAGGTG 480
QY      141 GlnAlaMetLeuGlnGlnSerThrGlnGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
      |||||||
Db      481 CAGGCCATGCTCGGCCACAGACACGAGACCTGGGGTGGCGCTCGCTCCACACTCGCC 540
QY      161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
      |||||||
Db      541 AAGCTGGCTAAGCGGCTCCCTCCGATGCCGATGACCTCGAAGAGCGCTGGCAGGTAC 600
QY      181 GlnAlaGlnValArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
      |||||||
Db      601 CAGCGCGGGCGCGCGAGGCGCGGAGCGCGGCTCAGCGCCATCCGAGCGCGCTGGGG 660
QY      201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
      |||||||
Db      661 CCCCTGTGGAGACAGGCGCGCGCTGGCGCGCGCACGTGGCTCCCTGGCGCGCGCAGCG 720
QY      221 LeuGlnGlnArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
      |||||||
Db      721 CTACAGAGCGGCGCGCGCTGGCGCGCGCTGGCGCGCGCTGGCGCGCGATGAGATGGC 780
QY      241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
      |||||||
Db      781 AGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
QY      261 LeuGlnGlnGlnAlaGlnGlnLysArgLeuGlnAlaGlnValAspMetGlnArgLeuLys 280
      |||||||
Db      841 CTGAGAGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY      281 SerTrpPheGlnProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
      |||||||
Db      901 AGCTGTTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
QY      301 ValGlnAlaAlaValGlnThrSerAlaAlaProValProSerAspHis 317
      |||||||
Db      961 GTGCAGGCTGCTGCGGACACGCGCGCGCTGTGTGCGCGCGCGCGCGCGCGCGCG 1011

```

RESULT 6
 BC003557 1186 bp mRNA linear PRI 12-JUL-2001
 LOCUS
 DEFINITION Homo sapiens, apolipoprotein E, clone MGC:1571 IMAGE:3355712, mRNA,
 complete cds.
 ACCESSION BC003557 GI:13097698
 VERSION BC003557.1
 KEYWORDS Homo sapiens.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1186)
 AUTHOR STRAUSBERG, R.
 TITLE Direct Submission
 JOURNAL Submitted (20-FEB-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 CONTACT Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@cgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,

Susanna Chan, Readman Chin, Chris Fjell, Erin Garland, Ran Guin, Letlicia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stolt, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>
Series: IRAL Plate: 6 Row: h Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4557324.

FEATURES

source

1..1186
/organism="Homo sapiens"
/db_xref="LocusID:348"
/db_xref="taxon:9606"
/clone="MGC:1571 IMAGE:3355712"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NH.MGC_16"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
46..999
/codon_start=1
/product="apolipoprotein E"
/protein_id="AAH03557.1"
/db_xref="GI:13097699"
/translation="MAKYMALLVTFPLAGCAQKVEQVETPEPELNOQTEWQGRM ELAIGRFMDYLRWVOTLSBOVELLSOVQDELALMDETMKLAKYKSLEROLP VAEETRARLSKEILAAQARLQADMEDVCGRLVQYRGVQAMLGSTELRYLASLR KLRRLRLADDDLOKRLAVYQAGAREGRLSIRREGLVYVGGVRAVATVSLG OPIQERAAQWGERLRARMEGSRTRDLEVRKQVAVAKLEQAOQIRLQAEARQ ARLSWEPFLPVEDRQWAGIKVEKVAAGVSAAPVPSDNH"

CDS

BASE COUNT 248 a 366 c 425 g 147 t
ORIGIN

Alignment Scores:

Pred. No.: 9.97e-81 Length: 1186
Score: 1585.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-827-854-15 (1-317) x BC003557 (1-1186)

QY 1 MetLysValLeuTPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 46 ATGAAGTTCTGTGGGCTGCGTGTGTCACATTCCTGGCAGATGCCAGGCCAAGTGT 105
QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnThrGlnSer 40
DB 106 GAGCAAGGGGTGGAGACAGCCGAGCCGAGCGCCGACACACAGCAGGCGCAGAGC 165
QY 41 GlyGlnArgTPGLeuAlaLeuGlnArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 166 GGGCAGCGCTGGAACTGGCAGCTGCTTTGGATTTACCTGCTGGGTGCAGACA 225
QY 61 LeuSerGlnGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 226 CTGTCTGAGCAGTGCAGAGAGAGCTGCTCCAGGTCCAGCAGCAGAACTGAGGCGC 285
QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
DB 286 CTGATGACGAGACCATGAGAGTTGAAGCCTTACAACTGGAACTGGAGAAACTGTG 345
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLysSerLysGlnLeuGlnAlaGlnAla 120
DB 346 ACCCGGTGGCGAGAGAACCGGGCAGCGCTGTCCAAGAGAGCTCAGGCGGCGAGGCC 405
QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlnArgLeuValGlnTyrArgGlnVal 140

DB 406 CGCGTGGCGCGACATGAGAGACGTGTGCGCGCTGTGTGCTACCGCGCGAGGTG 465
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 466 CAGGCAATGCTGGCCAGAGACACGAGAGCTGCGGTGCGCTGCTCCCACTGGCG 525
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 526 AAGCTGCTTAAGCGGCTCTCCGCGCATGTGACATGACAGACGCGCTGGCACTGTAC 585
QY 181 GlnAlaGlnAlaArgGlnGlnGlnAlaGlnArgLysSerAlaAlaArgGlnArgLeuGly 200
DB 586 CAGCGCGGGCGCGGAGAGCGCGCAGCGCGGCTCAGCGCCATCCGAGCGCGCTGGGG 645
QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnTyrPro 220
DB 646 CCGCTGGTGAACAGAGCGCGCTGCGGCGCGCTGTGTGCTGCTGCGCGCGCGCG 705
QY 221 LeuGlnGlnArgAlaGlnAlaAlaTrpGlnLysArgLeuArgAlaArgMetGlnGlnMetGly 240
DB 706 CTACAGAGAGCG 765
QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
DB 766 AGCGGAGCG 825
QY 261 LeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
DB 826 CTGGAGGAGAGCG 885
QY 281 SerTrpPheGlnProLeuValGlnLysPheGlnArgGlnTrpAlaGlnLeuValGlnLys 300
DB 886 AGCTGTTGAGACCGCTGTGTGAGAGACATGACGCCGACGTGGCGCGCGCTGGTGAAGA 945
QY 301 ValGlnAlaAlaValGlnThrSerAlaAlaProValProSerAspAsnHis 317
DB 946 GTGCAGCGCTGCGGTGGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 996

RESULT 7

E08423 1110 bp RNA linear PAT 29-SEP-1997
LOCUS E08423 DNA coding human apolipoprotein E3.
DEFINITION E08423
ACCESSION E08423.1 GI:2176540
VERSION E08423.1
KEYWORDS JP 1994J15392-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1110)
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Euteleostomi; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS Morimoto, H. and Teranishi, Y.
TITLE METHOD FOR PRODUCING APOLIPOPROTEIN
JOURNAL Patent: JP 1994J15392-A 1 15-NOV-1994;
MITSUBISHI KASEI CORP

COMMENT

OS Homo sapiens (human)
PN JP 1994J15392-A/1
PD 15-NOV-1994
PF 11-JUN-1985 JP 1994015433
PI MORIMOTO HIRONORI, TERANISHI YUTAKA
PC C12P21/02, C07K13/00, C12N5/10, C12N5/12, C12N5/18, C12P21/02,
PC C12R1/91,
PC (C12N5/10, C12R1/91):
CC strandedness: Double;
CC topology: Linear;
FH Key
FH Location/Qualifiers

FT source 1..1110
FT /organism="Homo sapiens"
FT sig_peptide 15..68
FT mat_peptide 69..968
FT /product="human apolipoprotein E3".
FEATURES
source Location/Qualifiers
1..1110


```

/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT      198 a      355 c      414 g      143 t
ORIGIN
Alignment Scores:
Pred. No.:      2.3e-80      Length:      1110
Score:          1578.00      Matches:      316
Percent Similarity: 99.68%      Conservative: 0
Best Local Similarity: 99.68%      Mismatches: 1
Query Match:      99.56%      Indels:      0
DB:              Gaps:      0

US-09-827-854-15 (1-317) x E08423 (1-1110)
QY      1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
      15 ATGAAGGTTCTGTGGGCTGTGCTGTGCTACATCTCTGGCAGAGATGCCAGGCCAAGGTC 74
QY      21 GUGGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
      75 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCGGCAGACAGCCGAGTGGCAGAGC 134
QY      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
      135 GGGCAGCGCTGGGAAGCTGGACCTGGGCTGCTTTGGATTACCTGGCTGGGTGGCAGACA 194
QY      61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
      195 CTGTCTGACAGAGTGCAGAGAGAGAGCTGTCTACGTCTCCAGAGTCCAGAGACTGAGAGGCC 254
QY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu 100
      255 CTGATGAGCAGACCATGAAGAGTTGAAGGCTTACAAATCGGAAGCTGGAGGAACAATG 314
QY      101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
      315 ACCCCGGTGGCGGAGAGACGGCGGACGCTGTCCAGAGAGCTGCAGCGCGGCGAGGCC 374
QY      121 ArgLeuGluValAlaAspMetGluAspValLysGlyArgLeuValGlnThrArgGlyGluVal 140
      375 CGGCTGGCGCGGACATGAGAGAGCTGTGCGGCGCTGTGTGATGATACCGCGCGAGAGTC 434
QY      141 GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
      435 CAGGCGATGCTGGCGCAGAGACCGAGAGCTCGGGTGGCGCTCCCTGCCCTCCAGCTGGCG 494
QY      161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
      495 AAGCTCGTGAAGCGGCTCTCCGCGATGCCATGACCTGCAGAAAGCGCTGGCAGATGTAC 554
QY      181 GlnAlaGluValAlaArgGluGluGluArgGlyLeuSerAlaThrLeuArgGluArgGly 200
      555 CAGGCGGCGCGCGGACAGGCGCGGCGGCGGCTTACAGCGCATCCGCGAGCGCTCGGGG 614
QY      201 ProLeuValGluGlnGluLysArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
      615 CCCCTGGTGAACAGGCGCGCGTGGCGGCGGCACATGTGGGCTCCCTGGCGCGGCGCGCG 674
QY      221 LeuGlnGluArgAlaGlnAlaThrGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
      675 CTACAGAGAGCGGGCCAGGCTGTGGGCGAGCGGCTGCGGCGGAGTGAAGAGATGGGCG 734
QY      241 SerArgTrpArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
      735 ACGCGGACCGCGGACCGCTGGGCGAGGTGAAGAGACAGAGTGGCCAGAGTGGCGCGCAAG 794
QY      261 LeuGluGluGlnAlaGlnGlnThrArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
      795 CTGGAGAGACAGAGCCAGAGATGAGCTTGCAGGCGGAGGCGCTCCAGGCGCGGCTCAAG 854
QY      281 SerTrpPheGluProLeuValGluLysPheGlnArgGlnTrpAlaGlyLeuValGluLys 300
      1110

```

```

DB      855 AGCTGTTGAGACCCCTGTGTGACAGACATGCACGCCCACTGGGCGGCTGTGGAGAAG 914
QY      301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
      915 GTGCAGGCTGGCGGTGGGCAACAGCGCGCGCTGTGCTCCAGGAGCAATCAG 965
RESULT 8
BD004277
LOCUS      BD004277      1156 bp      DNA      linear      PAT 31-JAN-2002
DEFINITION Apo E humanized mammal.
ACCESSION BD004277
VERSION    BD004277.1 GI:18632238
KEYWORDS   JP 2001017028-A/1.
SOURCE
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 1156)
AUTHORS    Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.
TITLE      Apo E humanized mammal
JOURNAL    Patent: JP 2001017028-A 1 23-JAN-2001;
            MITSUBISHI CHEMICAL CORP
COMMENT    OS Homo sapiens (human)
            PN JP 2001017028-A/1
            PD 23-JAN-2001
            PF 28-APR-2000 JP 2000128919
            PR
            PI SHINOBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA PC
            A01K67/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10, PC
            C12N15/09//C07K14/775,
            PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,(C12N5/00,C12R1:91) CC

FEATURES
source     1..1156      Location/Qualifiers
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT 208 a      367 c      432 g      149 t
ORIGIN
Alignment Scores:
Pred. No.:      2.72e-80      Length:      1156
Score:          1577.00      Matches:      316
Percent Similarity: 99.68%      Conservative: 0
Best Local Similarity: 99.68%      Mismatches: 1
Query Match:      99.50%      Indels:      0
DB:              Gaps:      0

US-09-827-854-15 (1-317) x BD004277 (1-1156)
QY      1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
      61 ATGAAGGTTCTGTGGGCTGTGCTGTGCTACATCTCTGGCAGAGATGCCAGGCCAAGGTC 120
QY      21 GUGGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
      121 GAGCAAGCGGTGGAGACAGAGCCGAGAGCTGCGGCGCAGACAGCCGAGTGGCAGAGC 180
QY      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
      181 GGGCAGCGCTGGGAAGCTGGACCTGCTTTGGATTACCTGCGCTGGGTGGCAGACA 240
QY      61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
      241 CTGTCTGACAGAGTGCAGAGAGAGCTGTCTACGTCTCCAGAGTCCAGAGACTGAGAGCG 300
QY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu 100
      301 CTGATGAGACAGACCATGAAGAGTTGAAGGCTTACAAATCGGAAGCTGAGAGACAATG 360
QY      101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120

```


Db 361 ACCCGGTGGCGAGAGACGGCGCACGGCTGTCCAAAGAGATGCAGGCGGCGAGCC 420
 Oy 121 ArgLeuGluAlaAspMetGluAspValIcysGluArgLeuValGlnTyrArgGluVal 140
 Db 421 CGCGTGGCGGCGGACATGAGACCTGTGCGGCCCTGGGTCAGTACGGCGGAGGTG 480
 Oy 141 GlnAlaMetLeuGluGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 481 CAGGCGATGCTGGCCACAGACAGCAGAGAGCTCGGGTGGCTCCCTCCACCTGGCC 540
 Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 Db 541 AAGCTGCTAAGCGGCTCTCCCGCATGCTCCGATGCTGCAAGATGCTGGCACTGAC 600
 Oy 181 GlnAlaGluAlaArgGluGluValGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db 601 CAGGCGGGGGCGCGGAGGGCGCGAGCGGCTCAGCGGCTCAGCGGAGCGCTGGGG 660
 Oy 201 ProLeuValGluGlnGluArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 661 CCCCTGTGGAAACAGAGGGCGCGTGGGGCCGCACTGTGGCTCCCTGGCGGCGAGCCG 720
 Oy 221 LeuGlnGluArgAlaGlnAlaTrrpLysLysLysLysLysLysLysLysLysLys 240
 Db 721 CTACAGAGCGGGCGCCAGGCTGGGGCGAGCGGCTGCGCGCGGATGAGAGATGGGC 780
 Oy 241 SerArgThrArgAspArgLeuAspLysLysLysLysLysLysLysLysLysLys 260
 Db 781 AGCGCGACCGCGACCGCTGGAGAGGTGAAGAGAGGTGGCGGAGGTGGCGGCCAG 840
 Oy 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleArgLeuLys 280
 Db 841 CTGGAGAGAGAGCGCCACAGATACGCTTCGAGCGCGAGGCTTCAGAGCCGCTCAAG 900
 Oy 281 SerTrrpHeGluProLeuValGluAspMetGlnArgGlnTrrpAlaGlyLeuValGluLys 300
 Db 901 AGCTGTTCGAGCGCCCTGTGGAAACATGCAGCGCCAGTGGCGGCTGGTGGAAG 960
 Oy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 961 GTGCAGGCTGCGGTGGSCACAGCGCGCCCTGTGTCCAGACAAATCAC 1011
 RESULT 9
 AX333278 1157 bp DNA linear PAT 09-JAN-2002
 LOCUS Sequence 3787 from Patent W00194629.
 ACCESSION AX333278
 VERSION AX333278.1 GI:18123912
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
 Horrikan,S., Soppet,D.R. and Weaver,Z.
 Cancer gene determination and therapeutic screening using signature
 TITLE
 JOURNAL Patent: WO 0194629-A 3787 13-DEC-2001;
 AVALON Pharmaceuticals (US)
 FEATURES
 source 1. 1157
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 212 a 370 c 426 g 149 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.1e-80 length: 1157
 Score: 1576.00 Matches: 315
 Percent Similarity: 99.37% Conservative: 0
 Best Local Similarity: 99.37% Mismatches: 2
 Query Match: 99.43% Indels: 0

DB: 6 Gaps: 0
 US-09-827-854-15 (1-317) x AX333278 (1-1157)
 Oy 1 MetLysValLeuTrrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 62 ATGAAGGTTCGTGGGCTGCGTGTGGTGCACATTCCTGGCAGAGATGCCAGGCCAAGGTG 121
 Oy 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrrpLys 40
 Db 122 GAGCAAGCGGTGGAGACAGAGCGGAGCGCGAGCGCGCCAGCAGACAGAGAGTGGCAGAG 181
 Oy 41 GlyGlnArgTrrpGluLeuAlaLeuGlyLysArgPheTrrpAspTrrpLysArgTrrpValGlnThr 60
 Db 182 GCGCCAGCGCTGGGAACCTGGCACTGGGTGGCTTTTGGGATTAATCTGCTGGGTGGCAGACA 241
 Oy 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 Db 242 CTGTCTGAGAGGTGCAGAGAGAGCTGCTCAGCTCCCAAGTACCCCAAGAACTGAGGCGC 301
 Oy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrrpLysSerGluLeuGlnGluLeu 100
 Db 302 CTGATGGACGAGACCATAGAGATTGAAGGCTTCAATCGGAATCGGAATCGGAACACTG 361
 Oy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 Db 362 ACCCGGTAGCGGAGAGAACCGCGGACGCGCTGTCCAAAGAGACTCAGAGCGCGAGGCC 421
 Oy 121 ArgLeuGlnAlaAspMetGluAspValIcysGluArgLeuValGlnTyrArgGluVal 140
 Db 422 CGCGTGGCGGCGGACATGAGAGAGCTGTGGCGGCCCTGTGTGACATACCGCGGAGGTG 481
 Oy 141 GlnAlaMetLeuGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 482 CAGGCGATGCTCGGCAAGACACCGAGAGGCTGGCGGCGGCTGCTCCACCTGGCGC 541
 Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 Db 542 AAGCTGCTAAGCGGCTCTCCCGCATGCCATGACTGCAGAGAGGCTGGCACTGAC 601
 Oy 181 GlnAlaGluAlaArgGluGluValGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db 602 CAGGCGGGGGCGCGGAGGGCGCGAGCGGCGGCTCAGCGGCTCAGCGGAGCGCTGGGG 661
 Oy 201 ProLeuValGluGlnGluArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 662 CCCCTGTGGAAACAGAGGGCGCGTGGGGCCGCACTGTGGCTCCCTGGCGGCGAGCCG 721
 Oy 221 LeuGlnGluArgAlaGlnAlaTrrpLysLysLysLysLysLysLysLysLysLys 240
 Db 722 CTACAGAGCGGGCGCCAGGCTGGGGCGAGCGGCTGGCGCGGATGGAGAGATGGGGC 781
 Oy 241 SerArgThrArgAspArgLeuAspLysLysLysLysLysLysLysLysLysLys 260
 Db 782 AGTCGAGACCGCGACCGCTGGAGAGGTGAAGAGAGGTGGCGGAGGTGGCGGCCAAG 841
 Oy 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleArgLeuLys 280
 Db 842 CTGGAGAGAGAGCGCCACAGATACGCTTCGAGCGCGAGGCTTCAGAGCCGCTCAAG 901
 Oy 281 SerTrrpHeGluProLeuValGluAspMetGlnArgGlnTrrpAlaGlyLeuValGluLys 300
 Db 902 AGCTGTTCGAGCGCCCTGTGGAAACATGCAGCGCCAGTGGCGGCTGGTGGAAG 961
 Oy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 962 GTGCAGGCTGCGGTGGSCACAGCGCGCCCTGTGTCCAGACAAATCAC 1012
 RESULT 10
 AX409597 1157 bp DNA linear PAT 14-JUN-2002
 LOCUS Sequence 2244 from Patent W00229103.
 ACCESSION AX409597

VERSION	AX409597.1	GI:21442302
KEYWORDS		
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
TITLE	Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.	
JOURNAL	Gene expression profiles in liver cancer	
GENE	Patent: WO 0229103-A 2244 11-APR-2002;	
LOGIC	GENE LOGIC INC (US)	
location/Qualifiers		
1..1157		
/organism="Homo sapiens"		
/db_xref="taxon:9606"		
/note="EMBL/Genbank Accession No. M12529"		
212 a	370 c	426 g 149 t
BASE COUNT		
ORIGIN		

Alignment Scores:	
Pred. No.:	3,1e-80
Score:	1576.00
Percent Similarity:	99.37%
Best local Similarity:	99.37%
Query Match:	99.43%
DB:	6
US-09-827-854-15 (1-317) x AX409597 (1-1157)	
	Length: 1157
	Matches: 315
	Conservative: 0
	Mismatches: 2
	Indels: 0
	Gaps: 0

QY	1	MethylsalicylauripalalaleuleulevalThrPheuleualaglycysinalalaval	20
Db	62	ATGAAGGTTCGTGGCGCTTCTGTGTCACTTCTGCGAGATCCAGGCCAAGTG	122
QY	21	GlucinalalavalglutrhgluprogluuprogluendargnglnthrtpgluSer	40
Db	122	GAGCAAGCGGTGGAGACAGAACCCGAGCCGCTGGCCAGCAGCCAGTGGCAAGC	181
QY	41	GlygluInArgTPgluLeuAlaIleuGlyArgPheTrpAspIyrLeuArgTPvalGlnThr	60
Db	182	GGCCAGCGCTGGAACTGGACTGTGGTGGCTTTTGGATTAACCTGGCCCTGGGTGCAGACA	241
QY	61	LeuSergluInvalglngluIngluLeuLeuSerSerGlnvalThrhngluLeuArgAla	80
Db	242	CTGTCTGACGAGTGGACGAGAGAGACTGTCTCAGCTCCCAAGTACCCAAAGATGAGGGCG	301
QY	81	LeuMetaspgluTrpMetIysgluLeuLysAlaIryIysSergluLeuIngluInleu	100
Db	302	CTGTATGACCGAGACCATTGAAGATTGAAGCCCTACAAATCTGGAACTGGAGGAACAACCTG	361
QY	101	ThrProValalagluguThrArgAlaArgPheuserIysgluLeuGlnAlaIleAlaGlnAla	120
Db	362	ACCCGGAGCGGAGAGAGACCGCGGGCAGCGCTGTCCAAAGAGAGCTGCAGAGAGCGCCAGGCC	421
QY	121	ArgLeuGlAlaAspMetgluAspValCysGlyArgLeuValGlnIrrArgIylgluVal	140
Db	422	CGCGTGGCGCGGACATGGAGAGAGTGTGGCGCGCTGTGTGAGATCCCGCGGAGGTG	481
QY	141	GlnIleMetIeuGlyInserThrGluIngluLeuAlaArgIleArgLeuAlaSerHisIeuArg	160
Db	482	CAGGCCATTGCTCGGCCAGAGACCCGAGAGAGTGGGGGTGGCGCTCGCTCCACCTGGCG	541
QY	161	LysLeuArgIysArgLeuLeuArgAspAlaAspAspIeuGlnIlysArgLeuAlaValTyr	180
Db	542	AAGCTGGTAAAGCGGCTCTCTCGGATCCGATGACCTGCAGAAAGCCCTGGCAGTGTAC	601
QY	181	GlnAlaGlyAlaArggluGlyAlaGluArgGlyLeuSerAlaIleArggluArgLeuGly	200
Db	602	CAGCGCGGGCGCCGAGGGGCGCGAGCGCGGCTTCAGCCCATCCCGAGCGACTGGGG	661
QY	201	ProleuValGluGlnGlyArgValArgAlaIleThrValGlySerIeuAlaGlyInPro	220
Db	662	CCCTCGATGAAACAGGGCCCGGTGGGGCCGACACTGTGGATCCCTGGCGCGGCGCAAGCG	721

Oy	221	LeuGluGluValArgAlaGlnAlaIleArgGluValIleArgLeuValIleArgMetGluGluMetGly	240
Db	722	CTACAGAGAGCGGGCCCAAGCCTGGGGGAGAGCGCTGGCGCCGCCGAGAGAGAGAGATGGC	781
Oy	241	SerArgThrArgAspArgLeuAspGluValIleGluGluGlnValAlaGluValArgAlaLys	260
Db	782	AGTGGAGCCCGCCGAGCCGCTGGAGAGAGTGAGAGCAGGTGGCGGAGTGGCCGCCAAG	841
Oy	261	LeuGluGluGluGlnAlaGlnGlnIleArgLeuGlnIleAlaGluLysIlePheGlnAlaArgLeuLys	280
Db	842	CTGGAGAGAGCGAGGCCCAAGATACGCTCTGAGGCCCAAGGCTTCCAGGCCCGGCTTCAG	901
Oy	281	SerTrpPheGluProLeuValGluAspMetLeuArgIleTrpAlaGlyLeuValGluLys	300
Db	902	AGCTGGTTCGAGCCCTGGTGTGAGACATCATCAGCCAGCATGGAGCCGGGCTGTGGAGAG	961
Oy	301	ValGlnAlaAlaValIleGlyThrSerAlaAlaProValIleProSerAspAsnHis	317
Db	962	GTCAGAGCTGCGGTGGGACCAAGCGCCCGCTGTGGCCAGACAAATCAC	1012

LOCUS	DEFINITION	ACCESSION	VERSION	RESULT 11
115975	Sequence 1 from patent US 5472858.	115975	115975.1	GI:1250883
				1157 bp DNA linear PAT 03-APR-1996

NETWORKS	.
SOURCE	Unknown.
ORGANISM	Unknown.

REFERENCE	1 (bases 1 to 1157)
AUTHORS	Attie,A.D., Gretch,D.G., Sturley,S.L. and Beckage,N.E
TITLE	Production of recombinant proteins in insect larvae
JOURNAL	Patent: US 5472858-A 1 05-DEC-1995;
FEATURES	Location/Qualifiers
SOURCE	1..1157

BASE COUNT	212 a	370 c	426 g	149 t
ORIGIN				

Alignment Scores:	
Pred. No.:	3.1e-80
Score:	1576.00
Percent Similarity:	99.37%
Best Local Similarity:	99.37%
Query Match:	99.43%
DB:	6
Length:	1157
Matches:	315
Conservative:	0
Mismatches:	237
Indels:	0
Gaps:	0

US-09-827-854-15 (1-317) x I15975 (1-1157)

OY	1	MeysValLeuTPPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIleVal	20
Db	62	ATGAAGGTCTGTGGGTGGCTTCTGTGTCAATCTCTGGCAGAGATCCAGGCCAAGTG	121
OY	21	GIuGlnAlaIaValGIuThrGIuProGIuProGIuLeuArgGlnGlnThrGIuTPGIuSer	40
Db	122	GAGCAACCGGTGGGAGACAGCCGGACCCGACCTGGCCAGAGACCGAGTGGCAGAGC	181
OY	41	GIyGlnArgTPGIuLeuAlaLeuGlyArgPheThrAspIyLeuArgTPValGIuThr	60
Db	182	GGCCAGCGCTGGGAACCTGGCACTGGGTGGCTTTTGGGATTACCTGGCTGGTGGCAGAC	241
OY	61	LeuSerGIuGlnValGlnGIuGlnLeuLeuSerSerGlnValThrGlnIuLeuArgAla	80
Db	242	CTGTCTAGGAGGTGCAGGAGGAGCTCTACGTCCCAATCTCACCCAAAGACAGAGGGC	301
OY	81	LeuMetAspGIuThrMetIySGluLeuIyAlaIyIySerGIuLeuGlnGIuGlnIleu	100
Db	302	CTGATGACCGACGACCATGAGAGAGATTGAAGCCTACAAATCGGAACCTGGAGGAACAAC	361
OY	101	ThrProValaGIuGIuThrArgAlaArgLeuSerIySGluLeuGlnAlaIaGlnAla	120
Db	362	ACCCCGGTAGCGAGGAGAGCGGGCCAGCGCTGTCCAAAGAGCTGCAGAGCGCGAGGCC	421

OY	121	ArgLeuGluAlaAspMetGlyAspValGlySerGlyArgLeuValGlnTyrArgGlyGluVal	140
DB	422	CGGCTGGGGGGGACATGAGACGTGTGGCGGCCCTGGTGTGACATACCGCGAGGTG	481
OY	141	GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValAlaGlyLeuAlaSerHisLeuArg	160
DB	482	CAGGCGCATGCTGGGACAGACACCGAGAGCTGGCGGCGCCCTGCTCCACCTGGCG	541
OY	161	LysLeuArgGlySerGlyLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
DB	542	AAGCTGCTAAGCGGCTCTCTCCGATCCCATGACCTGCAGAAAGCGCTGGCAGTGTAC	601
OY	181	GlnAlaGlyAlaArgGlyGlyAlaGlyGlyLeuArgGlyLeuSerAlaAlaLeuArgGlyLeuGly	200
DB	602	CAGGCGGGGGGGCGGAGGGGGCGGAGCGGCGCTGCACAGCCATCCGCGAGCGCTGGGG	661
OY	201	ProLeuValGlnGlnGlyArgValAlaAlaThrValGlySerLeuAlaGlyGlnPro	220
DB	662	CCCCGTGGTGAACAGAGCGCGCTGGGGCGCCCACTGTGGGCTCCCTGGCGCCAGCGG	721
OY	221	LeuGlnGluArgAlaGlnAlaAlaTyrPylGluArgLeuArgAlaArgMetGluGluMetGly	240
DB	722	CTACAGAGAGCGGCGCCAGCGCTGGGGCGAGCGGCTGCGCGCGGATGAGAGATGGGC	781
OY	241	SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys	260
DB	782	AGTCGGACCCCGACCGCTGGAGAGGTGAAGAGAGAGTGGGAGGTGCGCGCCAG	841
OY	261	LeuGlnGlnGlnAlaGlnGlnLeuArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys	280
DB	842	CTGGAGGAGAGAGCGCCACAGATAGCTGCAAGCGCGGAGGCTTCCAGGCGCCCTCAG	901
OY	281	SerTyrPheGluProLeuValGlnAspMetGlnArgGlnTyrPalaGlyLeuValGluLys	300
DB	902	AGCTGTTCGAGCCCTGTGTGAAGACATGACAGCGCCAGTGGCGCGGCTGTGTGAAG	961
OY	301	ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis	317
DB	962	GTCAGGCTGCGTGGGACAGCGCCGCTGGCCAGCAGCAATCAC	1012
RESULT 12			
HUMAPOE	LOCUS	HUMAPOE	1157 bp mRNA linear PRI 08-AUG-1995
DEFINITION	Human apolipoprotein E mRNA, complete cds.		
ACCESSION	M12529		
VERSION	M12529.1 GI:178848		
KEYWORDS	apolipoprotein.		
SOURCE	Homo sapiens (clone: pHAEl112,178,813.) male 57-year old liver CDNA to mRNA.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	McLean,J.W., Elshourbagy,N.A., Chang,D.J., Mahley,R.W. and Taylor,J.M.		
TITLE	Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing of a new variant		
JOURNAL	J. Biol. Chem. 259 (10), 6498-6504 (1984)		
MEDLINE	84212473		
PUBMED	6327862		
FEATURES	Location/Qualifiers		
source	1..1157		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/map="19q13.2"		
	/clone="pHAEl112,178,813."		
	/sex="male"		
	/tissue_type="liver"		
	/dev_stage="57-year old"		
	1..1157		
	/gene="APOE"		
	<1..1157		
MRNA			

DB	LOCUS	DEFINITION	APC E humanized mammal.	BD004279	1156 bp	DNA	linear	PAT 31-JAN-2002
DB	302	CHGATGACGACCATGATGAGAGATTGAAGGCCCTACAAATTCGAACTGGAGGAAACACTG	361					
OY	101	ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaGlnAla	120					
DB	362	ACCCGGGTAGCCGAGAGACGCGGGCACGCGCTGTCCAAAGACCTGCAGACGCGCAGCGC	421					
OY	121	ArgLeuGluAlaAspMetLysAspValLysGlyArgLeuValGlnThrArgLysGlyVal	140					
DB	422	CGGCTGGGCGCGACATGGAGACGtGtGCGCGCTGTGTACATGACGGCGGAGGtG	481					
OY	141	GlnAlaMetLeuGlyGlnSerThrGluLeuLeuArgValArgLeuAlaSerHisLeuArg	160					
DB	482	CAGGCATCTCTCGCCAGAGACCCAGAGAGCTGCGGGTCCGCTCCGCCACCTCGCC	541					
OY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180					
DB	542	AAGCTGCGTAAACGGCTCTCTCCGCAATCCCGATGACCTGCACAAAGCGCTGCACAGTAC	601					
OY	181	GlnAlaGlyAlaArgGluGluGluAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200					
DB	602	CAGGCGGGGGCCCCGGAGGGGCCCGAGCGCGGCTCAGCGCATCCGAGCGGCGGCGG	661					
OY	201	ProLeuValGluGlnGlyArgValArgAlaAlaIleThrValGlySerLeuAlaGlyLysPro	220					
DB	662	CCCTGTGTGGAACAGGGCCGCTGGGGCCGACCTGTGGCTCCCTGGCCGCGCAGCGC	721					
OY	221	LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGlnMetGly	240					
DB	722	CTACAGAGAGGGGCCAGGCCCTGGGGCCAGCGCGCTCCGCGCGCATGGAGAGATGGGC	781					
OY	241	SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys	260					
DB	782	AGTCGACCCCGGACCGCTCTGGACAGAGGGAAGAGAGAGTGGCGAGGTGGCGCCCAAG	841					
OY	261	LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys	280					
DB	842	CTGGAGAGACAGGCCACAGATACGCTCTGCAGGCGGAGCCCTTCAGGCCGCTCAAG	901					
OY	281	SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys	300					
DB	902	AGCTGGTTCGACCCCTGTGTGGAGACATGACAGCGCACTGGCGCGGCTGTGGAGAG	961					
OY	301	ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis	317					
DB	962	GTCAGAGCTGCGCTGGGACACAGCGCGCGCTGTGTCCAGGACACATTCAC	1012					
RESULT 13								
LOCUS	BD004279		1156 bp	DNA	linear	PAT 31-JAN-2002		
DEFINITION	APC E humanized mammal.							
ACCESSION	BD004279							
VERSION	BD004279.1							
KEYWORDS	JP 2001017028-A/3.							
SOURCE	Homo sapiens.							
ORGANISM	Homo sapiens							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.							
TITLE	I (bases 1 to 1156)							
JOURNAL	Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.							
COMMENT	Apoc E humanized mammal Patent: JP 2001017028-A 3 23-JAN-2001;							
	MITSUBISHI CHEMICAL CORP							
	OS Homo sapiens (human)							
	PN JP 2001017028-A/3							
	PD 23-JAN-2001							
	PF 28-APR-2000 JP 2000128919							
	PR							
	PI SHINOBU FUJITA,HIROKI HAMANAKA,YUKO FUKUI,MINESUKE YOKOYAMA							
	PC A0167/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10, C12N15/09//C07K14/775,							
	PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,(C12N5/00,C12R1:91) CC							

FEATURES	source	FT	CDS	Location/Qualifiers	Location/Qualifiers
				1..1156	/organism="Homo sapiens" /db_xref="taxon:9606"
BASE COUNT	208 a	369 c	432 g	147 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	4,56e-80	Length:	1156		
Score:	1573.00	Matches:	316		
Percent Similarity:	99.68%	Conservative:	0		
Best Local Similarity:	99.68%	Mismatches:	1		
Query Match:	99.24%	Indels:	0		
DB:	6	Gaps:	0		
US-09-827-854-15 (1-317) x BD004279 (1-1156)					
QY	1	MettysValleTrrpPaLaLaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal	20		
DB	61	ANGAAGTTCGTGTGGGCTGGCTTCTGTCTGCATTCCTCGGAGAGATCCAGGCCAAGTGC	120		
QY	21	GIUGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnThrGluTrrpGlnSer	40		
DB	121	GAGCAAGCGGTGGACACAGACCAGGCCACGCTGCCAGCAGACGAGAGTGGCAGAGC	180		
QY	41	GLYAlaGTrTPGluLeuAlaLeuGlyAraPheTrpAspTyrLeuAqTrpValGlnThr	60		
DB	181	GCCACGCGCTGGGAAGTGGCACTGGGTGCTTTTGGATTAACCTGGCGTGGGTGCAGACA	240		
QY	61	LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80		
DB	241	CTGTCTGACAGGTGCACAGAGAGACTGCTCAGCTCCAGGTACCCAGAACTGAGGCGC	300		
QY	81	LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu	100		
DB	301	CTGAATGGACGAGACCATGAAGAGTTGAAGGCTTCAAAATGGAACTGGAGAAACAATG	360		
QY	101	ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120		
DB	361	ACCCGGTGGGGAAGAGACCGCGGGCAAGCTGTGCCAAGGAGCTGCAGCGCGGCAGGCC	420		
QY	121	ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal	140		
DB	421	CGGCTGGGCGCGGACATGAGAGAGCTGCGCGCGCTGTGGTGCAGTACCAGCGAGAGTG	480		
QY	141	GlnAlaMetLeuGlyGlnSerThrGluLeuLeuArgValArgLeuAlaSerHisLeuArg	160		
DB	481	CAGGCGATGCTCGGCGACAGACACCGAGAGAGCTGGGGTGGCGCTCCCACTGCGC	540		
QY	161	LysLeuAraGlyArgLeuLeuAraGAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180		
DB	541	AAGCTGCGTAAAGCGGCTCTCCGGAATGCCGATACCTCGAGAAAGCCCTCGACAGTAC	600		
QY	181	GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200		
DB	601	CAGGCGGGGGCCCGGAGAGGGCGGAGCGGCTTCAAGCCCATCCGAGAGCCCTGGGG	660		
QY	201	ProLeuValGluGlnGlyArgValArgAlaAlaIleThrValGlySerLeuAlaGlyGlnPro	220		
DB	661	CCCGTGGTGAACAGGGCGCGGTGCGGGCGGCCCATCTGTGGGTCCTCGGCGGCAGCCG	720		
QY	221	LeuGlnGluArgAlaGlnAlaTrrpGlyGluArgLeuArgAlaArgMetGluGluMetGly	240		
DB	721	CTACAGGAGCGGGCCACGAGCTGGGGGAGAGGCTGGCGCGGAGAGGAGAGTGGGC	780		
QY	241	SerArgTrrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaIys	260		
DB	781	AGCCGAGCCCGCGAGCGCTGGAGAGAGTGAAGAGACAGGTGGCGAGGTGGCGGCCAG	840		
QY	261	LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys	280		

Db 841 CTGGAGAGCAGCCAGCAGATACGCTGACAGCCGCTTCCAGGCCCTCAAG 900
QY 281 SerTrrPheGluProLeuValGluAspMetGlnArgIntPrAlaGlyLeuValGluIys 300
Db 901 AACTGTTGACCCCTGGTGGAGACATGACGCCCATGTGGCCGCGCTGGTGGAGAAG 960
QY 301 ValGlnAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCGGTGGGACACGACGCCGCTGTGGCCAGCAGCAATCAC 1011
RESULT 14
MEAPOE 1178 bp mRNA linear PRI 31-MAR-1995
LOCUS DEFINITION Monkey mRNA for apolipoprotein E.
X13887
VERSION X13887.1 GI:38054
KEYWORDS apolipoprotein; apolipoprotein E.
SOURCE Macaca fascicularis.
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopitheciinae; Macaca.
REFERENCE 1 (bases 1 to 1178)
AUTHORS Marotti,K.R.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1988) Marotti K.R., The Upjohn Co, 7242 209 713,
301 Henrietta Street, Kalamazoo, MI 49008
2 (bases 1 to 1178)
AUTHORS Marotti,K.R., Whitted,B.E., Castle,C.K., Polites,H.G. and
Metchlor,G.W.
TITLE Nucleotide sequence of the cynomolgus monkey apolipoprotein E cDNA
JOURNAL Nucleic Acids Res. 17 (4), 1778 (1989)
MEDLINE 89160349
PUBMED 2922300
COMMENT Data kindly reviewed (20-Mar-1989) by Marotti K.R.
FEATURES
source
Location/Qualifiers
1..1178
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="1"
/tissue_type="liver"
/clone_lib="lambda gll1"
83..1036
/codon_start=1
/product="proapolipoprotein E (AA -18 to 299)"
/protein_id="CAA32092.1"
/db_xref="GI:38054"
/db_xref="SWISS-PROT:P10517"
/translation="MKYLMALVTFPLAGCOAKYEQPEPEPELROQAEQSGQPA
ELALGRFMDYLRWYOTLSEVOEELSPVOYOEITLTMDETMKELKAYKSLLEQLSP
VAEETPARSLKEIQAAQARLIGADMEQVSRVYORSEYQAMLGSTETELRLARLSHR
KLKRLRLDADLDQRLAVYQAGRBGAEKRVSAIREFLGLVDEQGRVRAATVGLAS
QPLDERAQLLESLRLARMEKMSRRLRDLDEKQVAEVRARLKEEQADQISLQAEARQ
ARLSWEPELVEDMRQWAGLYEKVQAAVASTAPVPIDNH"
83..1036
sig_peptide
mat_peptide
misc_feature
/product="apolipoprotein E (AA 1 - 299)"
1154..1159
/note="POLYA signal"
BASE COUNT 217 a 373 c 426 g 162 t
ORIGIN
Alignment Scores:
Pred. No.: 3,13e-73 Length: 1178
Score: 1451.00 Matches: 296
Percent Similarity: 94.01% Conservative: 2
Best Local Similarity: 93.38% Mismatches: 19
Query Match: 91.55% Indels: 0
DB: 9 Gaps: 0
US-09-827-854-15 (1-317) x MEAPOE (1-1178)
QY 1 MetLysValIleuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20

Db 83 ATGAAGTTCTGTGGGCTGGCTTGGTGGTGCACATTCCTGGAGATGCGCAGCCCAAGTG 142
QY 21 GluGlnAlaValAlaGlyThrGluProGluProGluLeuArgGlnGluThrGluTrpGlnSer 40
Db 143 GAGCAACCGGTGGAGCCAGACAGACGAGACCGAGCTTGGCAGACAGCTGAGGCGCAAGC 202
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTyrValGlnThr 60
Db 203 GGCAGACCCCTGGAGACTGGCACAATGGGTGCTTTTGGAGTAACTGCGCTGGGTGAGACA 262
QY 61 LeuSerGluGlnValAlaGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 263 CTGTCTGAGCAGCTGACAGAGAGAGCTGTACAGCCGCCAGGTCACCCAGCAACTGACAGACG 322
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluLeuGln 100
Db 323 CTGATGGAGACCAATGAAAGGTTGAAGGCCCTCAATTCGAACTGGAGAGACAGCTG 382
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 383 AGCCCGGTGGCGAGAGACGCGGCGACGCTGTCCAGAGACTGCAGCGCGCCAGGCC 442
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 443 CGGTGGGTGGCGACATGAGAGACGCGCGACCGCGCTGTGTCAGTACCGCAGCGAGGTG 502
QY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 503 CAGGCCATGCTGTGGCAGAGTACCAGAGCTGCGGGCGCGCGCTGCCCTCCACCTGTGCC 562
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 563 AAGCTGCCGACAGCGGCTCTCCCGCATGCTATGACTTCACAGAGCGCTGGCAGTGTAT 622
QY 181 GlnAlaGlyAlaArgGluGluValAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 623 CAGGCGGGGGCGCGGAGGCGCGGAGCGCGGCGGCGGATCCGAGCGCGCGCTGGGA 682
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 683 CCCCTGGGTGGACAGAGGCGCGCTGGCGCGCGCTGTGGCTCCCTGGCCAGCAGCGG 742
QY 221 LeuGlnGluArgAlaGlnAlaThrPglGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
Db 743 CTTCAGAGGCGGCGCCAGGCGCTTGGGTGAGCGGCTTCGGCGCGAGTGGAGAGATGGGC 802
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValAlaArgAlaIys 260
Db 803 AGCGGAGCCCGGAGCGCGCTGTGACAGAGGAGAGAGAGAGTGGCGAGGTGGCGCCCAAG 862
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 863 CTGGAGGAAACAGCCAGACAGATAAGCTGTGAGCGCGGAGGCGCTTCAGGCGCGCTCAAG 922
QY 281 SerTrrPheGluProLeuValGluAspMetGlnArgGlnIntPrAlaGlyLeuValGluIys 300
Db 923 AACTGTTGACCCCTGGTGGAGATATGACGCCCGATGGGTGGGTGGGTGGAGAGAG 982
QY 301 ValGlnAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 983 GTGCAGGCTGCGGTGGGCGCCAGCAGCGGCCCTGTGGCCATGACCAATCAC 1033
RESULT 15
AF261279 5491 bp DNA linear PRI 27-OCT-2000
LOCUS DEFINITION Homo sapiens apolipoprotein-E gene, complete cds.
ACCESSION AF261279
VERSION AF261279.1 GI:11034800
KEYWORDS Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	Mammalia: Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 5491) Nickerson,D.A., Taylor,S.L., Fullerton,S.M., Weiss,K.M., Clark,A.G., Stengard,J.H., Salomaa,V., Boerwinkle,E. and Sing,C.F.
TITLE	Sequence diversity and large-scale typing of SNPs in the human apolipoprotein E gene
JOURNAL	Genome Res. 10 (10), 1532-1545 (2000)
MEDLINE	20499366
PUBMED	11042151
REFERENCE	2 (bases 1 to 5491) Nickerson,D.A.
AUTHORS	Direct Submission
TITLE	Submitted (27-APR-2000) Department of Molecular Biotechnology, University of Washington, Box 357730, Seattle, WA 98195, USA
JOURNAL	Location/Qualifiers
FEATURES	1..5491 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="19" /map="19q13.2" /3..>108 /note="putative" /rpt_family="MIR" /rpt_type=dispersed 73 /frequency="0.01" /replace="t" /note=">295 /note="putative" /rpt_family="MIR" /rpt_type=dispersed 308 /frequency="0.01" /replace="t" /note=">339 /note="putative" /rpt_type=tandem /note=">637 /note="putative" /rpt_family="Alu" /rpt_type=dispersed 471 /frequency="0.01" /replace="g" 545 /frequency="0.01" /replace="t" 560 /frequency="0.22" /replace="t" 624 /frequency="0.07" /replace="c" /note=">718 /note="putative" /rpt_type=tandem 832 /frequency="0.45" /replace="t" /product="apolipoprotein-E" 1163 /frequency="0.35" /replace="c" 1522 /frequency="0.01" /replace="a" 1575 /frequency="0.01" /replace="t" /note="APOE" /codon_start=1

[illegible]

Alignment Scores:

Pseq. No.: 1.73e-69 Length: 5491
 Score: 1396.50 Matches: 304
 Percent Similarity: 60.80% Conservative: 0
 Best Local Similarity: 60.80% Mismatches: 2
 Query Match: 88.11% Indels: 194
 DB: 9 Gaps: 1

us-09-827-854-15 (1-317) x AF261279 (1-5491)

QY 12 PheLeuAlaGlyCysGlnAlaLysValGluGlnAlaValGluThrGluProGluProGlu 31
 DB 3003 TTCCACACAGATGCGCAGGCCAAGGTGGAGCAAGCGGTGGAGACAGAGCCGAGCCGAG 3062
 QY 32 LeuArgGlnGlnThrGluTrpGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPhe 51
 DB 3063 CTGCCCGACGACGAGTGGAGCGGCGCCAGCGCTGGGAACCTGGCACTGGGTGCTTT 3122
 QY 52 TrpAspTyrLeuArgTyrValGlnThrLeuSerGluGlnValGlnGluGluLeuLeuSer 71
 DB 3123 TGGGATTACCTGCGCTGGGTGAGACACTGCTGTGAGCAGGTGCAGAGAGAGCTGCTCAGC 3182
 QY 72 SerGlnValThrGlnGluLeu----- 78
 DB 3183 TCCGAGTCCACCGACGAGAACTGAGGTGTCCTCCATCTGGCCCTTGACCTCCTGGTG 3242
 QY 78 ----- 78
 DB 3243 GCGGCTATACCTCCCGACGTCAGGTTTCATTCTGCCCCGTGCGTAAGTCTGGGGGG 3302
 QY 78 ----- 78
 DB 3303 COTGGGTCTGCTGCTGCTCTAGCTTCCCTTCCCATTTCTGACTCTGCTTACTCTCTC 3362
 QY 78 ----- 78
 DB 3363 TGAATTCCTCTCTCAGCTTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCACT 3422
 QY 78 ----- 78
 DB 3423 GCTCTGCT 3482
 QY 78 ----- 78
 DB 3483 CTCACCTGTTGCCCCAGGCTGCTTGAACCTTCTGGGCTCAAGCATCTCCGCTGG 3542
 QY 78 ----- 78
 DB 3543 COTCCCAAGTGTGGATTAGAGCATGAGCACCTTGGCGGCTCTAGCTCTCTCT 3602
 QY 78 ----- 78
 DB 3603 TCGTCTGCTCTGCTCTGCTCTGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3662
 QY 78 ----- 78
 DB 3663 GCCTCTGCCCCCT 3722
 QY 78 ----- 78
 DB 3723 CCCCATTCCAGCCCTTCTCCCGCTCTCCACTGTGCGACACCTCCGCTCTGCGCG 3782
 QY 79 -ArgAlaLeuMetAspGluThrMetLeuLysGluLeuLysAlaTyrLysSerGluLeuGlu 98
 DB 3783 CAGGGCTGTGATGACGAGACCATGATGAGAGCTTGAAGGCTTACAAATCGAAGCTGGAGGA 3842
 QY 98 uGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAla 118
 DB 3843 ACAACTGACCCCGGTGGGAGAGACGCGGCGACGCGCTGTCCAAAGAGAGCTGCAAGCGCGC 3902
 QY 118 aGlnAlaArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGly 138
 DB 118 aGlnAlaArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGly 138

DB 3903 GCAGGCCGCTGGCGCGGACATGAGACGCTGTGCGCGCCCTGTGTGACTACCGCG 3962
 QY 128 YGluValGlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHis 158
 DB 3963 CGAGGTGCAAGGCTATGCTCTGGCGCAGACACGAGAGAGCTGCCGGTGTGCTCTGCTCCCA 4022
 QY 158 sLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAla 178
 DB 4023 CCGCGCAAGCGCTGTAGCGGCTCTCTCGCATGACCTGACGAGAACGCGCTGCG 4082
 QY 178 aValTyrGlnAlaGlyAlaArgGluGlyAlaGluArgLysSerAlaLeuArgLys 198
 DB 4083 AGTATACAGCGCGCGCGCGCGAGCGCGCGCGCGCGCTCACAGCCATCCGCGAGCG 4142
 QY 198 gLeuGlyProLeuValGluGlnGlyArgValArgAlaArgValAlaArgLysSerLeuAlaGly 218
 DB 4143 CCGGGGCCCCGTGTGGAACAGAGCGCGCGTGGCGCGCCCACTGTGGCTCTCTGCGCG 4202
 QY 218 YGlnProLeuGlnGluArgAlaGlnAlaArgLysGluArgLeuArgAlaArgMetGluGly 238
 DB 4203 CCAGCGCTACAGAGAGCG 4262
 QY 238 uMetGlySerArgTyrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArg 258
 DB 4263 GATGGCGAGCGGACCGCGACCGCTGTGAGAGGTGAAGAGCGAGTGGCGGAGGTGCG 4322
 QY 258 gAlaLysLeuGluGluGlnAlaGlnGlnLeuArgLeuGlnAlaGlnAlaArgLeuAlaArg 278
 DB 4323 CGCCAGGCTGGAGAGAGCG 4382
 QY 278 gLeuLysSerTyrPheGluProLeuValGluAspMetGlnArgGlnTyrPalaGlyLeuVal 298
 DB 4383 CCTCAAGAGCTGTGTGAGCGCGCTGTGTGAAGACATGACAGCGCGCGCGCGCGCGCGGT 4442
 QY 298 IGlulysValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 4443 GGAGAGGTGAGAGCTGCGGTGGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4500

Search completed: March 14, 2003, 17:32:12
 Job time : 2314.23 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:08:57 ; Search time 2307.23 Seconds

(without alignments)
3998.551 Million cell updates/sec

Title: US-09-827-854-16

Perfect score: 1589

Sequence: 1 MKVLMAALLVFLAGCOAKV.....VEKVOAVGTSAPVPSDNH 317

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+lp2n.model -DEV=xlp
-O=/cgn2.1/USPTO_spool/US09827854/runat_11032003_101610_27486/app-query.fasta.1.3576
-DB=genEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=-1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854.ecgn.1.1.13008.@runat_11032003_101610_27486 -NCPU=6 -ICPU=3
-NO_XLIPX -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6 -FCGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: GenEmbl:*
2: gb_ba:*
3: gb_hcg:*
4: gb_in:*
5: gb_com:*
6: gb_ov:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_com:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hlgo_hum:*
40: em_hlgo_mus:*
41: em_hlgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1589	100.0	1156	6	BD004277	BD004277 Apo E hum
2	1577	99.2	1110	6	E00359	E00359 CDNA coding
3	1577	99.2	1110	6	E00823	E00823 DNA sequenc
4	1577	99.2	1147	6	AX302545	AX302545 Sequence
5	1577	99.2	1156	6	BD004278	BD004278 Apo E hum
6	1577	99.2	1156	6	H004278	H004278 Homo sapien
7	1577	99.2	1186	9	BC003557	BC003557 Homo sapi
8	1570	98.8	1110	6	E08423	E08423 DNA coding
9	1568	98.7	1157	6	AX333278	AX333278 Sequence
10	1568	98.7	1157	6	AX409597	AX409597 Sequence
11	1568	98.7	1157	6	I15975	I15975 Sequence 1
12	1568	98.7	1157	6	H004279	H004279 Homo sapi
13	1565	98.5	1176	6	BD004279	BD004279 Apo E hum
14	1443	90.8	1158	9	MFAP0E	MFAP0E Monkey mRNA
15	1388.5	87.4	5491	6	AX58722	AX58722 Sequence
16	1388.5	87.4	41907	9	AF050154	AF050154 Homo sapi
17	1388.5	87.4	41907	9	AF050154	AF050154 Homo sapi
18	1388.5	87.4	107567	9	AC011481	AC011481 Homo sapi
19	1371	86.3	5515	9	H004278	H004278 Homo sapi
20	1342	84.5	5413	9	AF261280	AF261280 Pan trogl
21	1359	79.2	4762	9	BARAP0E	BARAP0E Baboon apol
22	1250	79.2	208239	2	AC021988	AC021988 Homo sapi
23	1250	78.7	1138	4	AF030380	AF030380 Tupai gl
24	1162	73.1	1060	4	RABAP0E	RABAP0E Rabbid apol
25	1155.5	72.7	965	6	AX384545	AX384545 Sequence
26	1155.5	72.7	1108	4	BTAP0LPE	BTAP0LPE Sequence
27	1155.5	72.7	5617	6	AX384541	AX384541 Sequence
28	1155.5	72.7	6026	6	AX384539	AX384539 Sequence
29	1143	71.9	718	9	AF20049753	AF20049753 Pan trogl
30	1141	71.8	1045	10	M0SAPOE	M0SAPOE Mouse apoli
31	1141	71.8	1104	10	BC028816	BC028816 Mus muscu
32	1139	71.7	718	9	AF20050052	AF20050052 Gorilla g
33	1139	71.7	718	9	AF20050053	AF20050053 Pongo pyg
34	1137.5	71.6	1154	4	BRAP0EMR	BRAP0EMR Baboon apol
35	1135	71.4	718	9	AF20050653	AF20050653 Mus muscu
36	1135	71.4	1122	4	SSAP0E	SSAP0E S.scrota mr
37	1130	71.1	1126	6	AR164342	AR164342 Sequence
38	1130	71.1	1126	6	AR205885	AR205885 Sequence
39	1086	68.3	951	10	S76779	S76779 RABAP0E
40	1064	67.0	959	10	M0SAPOE	M0SAPOE Mus muscu
41	1048.5	66.0	1069	10	RATAP0E	RATAP0E Rat apoli
42	1021	64.3	228698	2	AC127479	AC127479 Mus muscu
43	1021	64.3	237653	2	AC073760	AC073760 Mus muscu
44	995	62.6	4856	10	M0SAPE	M0SAPE Mus muscu
45	994	62.6	4267	6	AR164387	AR164387 Sequence

RESULT 1

ALIGNMENTS

BD004277	BD004277	1156 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	BD004277				
DEFINITION	Apo E humanized mammal.				
ACCESSION	BD004277				
VERSION	BD004277.1	GI:18632238			
KEYWORDS	JP 2001017028-A/1.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Primates; Carnivora; Hominoidea; Homo.				
AUTHORS	Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.				
TITLE	Apo E humanized mammal				
JOURNAL	Patent: JP 2001017028-A 1 23-JAN-2001;				
COMMENT	MITUBISHI CHEMICAL CORP OS Homo sapiens (human) PN JP 2001017028-A/1 PD 23-JAN-2001 PF 28-APR-2000 JP 2000128919 PR				
FEATURES	FT SHINOBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA PC source A01667/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10, PC C12N15/09/C07K14/775, PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,C12N5/00,C12R1:91) CC				
BASE COUNT	208 a 367 c 432 g 149 t				
ORIGIN					
Alignment Scores:					
Pred. No.:	2,66-83	Length:	1156		
Score:	1589.00	Matches:	317		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
US-09-827-854-16 (1-317) x BD004277 (1-1156)					
QY 1	MettysValleuTPPALaAlaLeuDeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20			
Db 61	ATGAAGCTTCGTGGGCTGCCTGCTGTCATCTCGCGAGATGCGCAGGCAAGGTG	120			
QY 21	GlutGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer	40			
Db 121	GAGCAAGGGGGGAGACACAGCCCGGACCGCAGCTGCGCCAGCAGACCGAGTGGCAGAGC	180			
QY 41	GlyGlnArgTPGluLeuAlaLeuGlyValArgPheTrpAspPyrLeuArgTrpValGlnThr	60			
Db 181	GGCCAGCCCTGGGAACCTGGCACCTGGGTGCTTTTGGATTACCTGCGCTGGGTGCACCA	240			
QY 61	LeuSerGlnGlnValGlnGlnGluLeuLeuSerSerGlnValTrpGlnGlnLeuArgAla	80			
Db 241	CTGTCTGACGACGAGTGCAGAGAGAGCTGCTCAGCTCCCGACGTCACCCAGAACTGAGGCG	300			
QY 81	LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu	100			
Db 301	CTGATGCGACGAGACCATGAAGAGATTGAAGGCTTAACAATCGGAATCGAGAACACTG	360			
QY 101	ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120			
Db 361	ACCCGGGGGGGAGAGACACGGGGCAGGGCTGTCCAAAGAGATGACAGCGGGCAGAGCC	420			
QY 121	ArgLeuGlnAlaAspMetGluAspValCysGlyThrArgLeuValGlnTyrArgGlyVal	140			
Db 421	CGGCTGGGGCGGACATGGAGAGACGTGTGGCGCCCGCTGGTGGCAGTACCGCGGCGAGGTG	480			
QY 141	GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg	160			

FEATURES	Location/Qualifiers
Db	481 CAGGCCATCTCGGCCACAGACACCGAGAGCTGCGGGTGGCTCGCTCCACTCCGCGC 540
Qy	161 LysLeuAaIyLysAArgLeuLeuAArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db	541 AAGCTCGTAAGGAGGCTCTCCGCGATGCCGATGACCTGCAGAAAGTGGCTGGCAGTGTAC 600
Qy	181 GlnAlaGlyAlaAArgLugLysAlaGluAArgGlyLeuSerAlaIleAArgLysAArgLugLys 200
Db	601 CAGGCGCGGGGCGCGCAGAGGGCCCGAGCGCGGCTTCAGGCCATCCGACGAGCGCTGGGG 660
Qy	201 ProLeuValGluGlnGlnGlyAArgValAArgAlaIleAArgValGlySerLeuAlaGlnPro 220
Db	661 CCCCTGTGTGAACAGGGGCGCGTGGCGGGCGCCACTGTGGGCTCCCTGGCCGGCCAGCGG 720
Qy	221 LeuGlnGluAArgAlaGlnAlaIleAArgLysAlaAArgLysAArgLysAArgLys 240
Db	721 CTACAGAGAGCGGGGCCAGGCGCTGGGGCGAGCGGCTGCGCGCGGATGTAGAGAGATGGGC 780
Qy	241 SerAArgThrAArgAspAArgLysAspGluValLysGluGlnValAlaGluValAArgAlaLys 260
Db	781 AACCGGACCGCGGAGCGCGCTGTGACGAGGTGAAGAGACAGTGGCGAGGTGGCGCGCAAG 840
Qy	261 LeuGlnGluGlnAlaGlnGlnIleAArgLeuGlnAlaGluAlaAArgAlaAArgLys 280
Db	841 CTGGAGAGACAGAGCCCGACAGATACGCTCTGCAGGCGCGAGGCTTCAGGCGCGGCTCAAG 900
Qy	281 SerTTPheGluProLeuValGluAspMetGlnAArgGlnTTPAlaGlyLeuValGluLys 300
Db	901 AACTGGTTTGAGACCCCTGTGTGGAAGCATGACAGCGCGCACTGGGCGGGCTGTGGAGAG 960
Qy	301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db	961 GTGCAGAGCTGCGGTGGGACACAGCGCGCGCTGTGGCCAGCAGACATTCAC 1011
RESULT 2	
E00359	
LOCUS	E00359 1110 bp RNA linear PAT 29-SEP-1997
DEFINITION	CDNA coding human apolipoprotein E3.
ACCESSION	E00359.1 GI:2168646
VERSION	JP 1985118189-A/1.
KEYWORDS	Homo sapiens.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens.
REFERENCE	Numakura, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1110) Teranishi, Y., Takamatsu, N., Matsui, Y., Kimura, M. and Ikeda, Y. DNA FRAGMENT Patient: JP 1985118189-A 1 25-JUN-1985; MITSUBISHI CHEM IND LTD
AUTHORS	OS human
TITLE	PN JP 1985118189-A/1
JOURNAL	PD 25-JUN-1985
COMMENT	PT 29-NOV-1983 JP 1983224980 PI TERANISHI YUTAKA, TAKAMATSU NOBUHIKO, MATSUI YASUSHI, PI KIMURA MASAKO. PI IKEDA YASUOKO PC C12N15/00, C07H21/04//C12P21/00; CC strandedness: Double; CC topology: Linear; CC hypothetical: No; CC anti-sense: No; CC *source: tissue=liverr; FH key Location/Qualifiers FH FH CDS 15..968 FT /product='apolipoprotein E3 precursor' FT FT sig_peptide 15..68 FT /product='apolipoprotein E3 signal peptide' FT FT mat_peptide 69..965 FT /product='apolipoprotein E3'. FT Location/Qualifiers


```

source 1..1110
/db_xref="taxon:9606"
BASE COUNT 198 a 353 c 416 g 143 t
ORIGIN

Alignment Scores:
Pred. No.: 1,22e-82 Length: 1110
Score: 1577.00 Matches: 316
Percent Similarity: 99.68% Conservative: 0
Best Local Similarity: 99.68% Mismatches: 1
Query Match: 99.24% Indels: 0
DB: Gaps: 0

US-09-827-854-16 (1-317) x E00359 (1-1110)
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 15 ATGAAGGTTCTGGGCTCGTTCGTGCTGACATTCCTGGCAGGATGCCAGGCCAAGGTG 74
QY 21 GUGGAlaValAlaGlnThrGluProGluLeuArgGlnGlnThrGlnThrGlnThr 40
DB 75 GAGCAAGCGGTGGAGACAGCCGGAGCCGAGCTGCCAGCAGACCGAGTGGCAGACC 134
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 135 GGCACAGCGCTGGGAACCTGGCAGCTGGCTCTTTGGGATTACCTGGCGTGGCAGACA 194
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 195 CTGTCTGACAGGTGACAGAGAGAGCTGCTCAGCTCCAGATCCAGACCCAGCACTGAGGGCG 254
QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnGlnLeu 100
DB 255 CTGATGAGACAGACCATGAGAGAGTTGAGAGCTTACAAATCGGACTGGAGAGACACTG 314
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
DB 315 ACCCGCGTGGCGAGAGACCGCGGCGCTGTCCAGAGAGCTGCAGCGCGGCGAGGCC 374
QY 121 ArgLeuGlyAlaAspMetLysAspValLysGlyArgLeuValGlnThrArgGlyGlnVal 140
DB 375 CGGCTGGGCGGACATGAGAGACGTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 434
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 435 CAGGCGATGCTGGCGCAGAGACACCGAGAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTG 494
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
DB 495 AAGCTCGTGAAGGGCTCTCCGCGATGGCGATGACCTGCAGAGAGCGCTGCGAGGTAC 554
QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 555 CAGGCGGGGCGCGCAGAGGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 614
QY 201 ProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
DB 615 CCCCTGTGGAAACAGGCGCGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 674
QY 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
DB 675 CTACAGAGACGGGCGCAGGCTGGGGCGAGCGGCTGCGCGCGGATGGAGAGAGAGGGG 734
QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
DB 735 AGCGGAGCCCGGACCGCTGGCGAGGTGAAGAGACAGTGGCGGAGGTGGCGGCGCAAG 794
QY 261 LeuGlnGlnGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
DB 795 CTGGAGGAGACGAGCCACAGATACGCTGCGAGGCGGAGGCTTCCAGGCGCGCTCAAG 854
QY 281 SerTrpPheGlnProLeuValGlnLysPheMetGlnArgGlnTrpAlaGlnLeuValGlnLys 300

```

```

DB 855 ACTGGTTGAGCCCTGCTGGAGACATGACAGCCGACCTGGGCGCGCTGGTGAGAG 914
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspHis 317
DB 915 GTGCAGGCTGCCGTGGGACAGCCGCGCGCTGCTGCCAGCAGCATAC 965

RESULT 3
E00823
LOCUS E00823
DEFINITION DNA sequence coding for human apolipoprotein E and its signal
peptide.
ACCESSION E00823
VERSION E00823.1 GI:2169084
KEYWORDS JP 1986096997-A/1.
SOURCE
ORGANISM
unclassified.
REFERENCE
1 (bases 1 to 1110)
AUTHORS Teranishi, Y., Matsui, Y., Ikeda, Y. and Kimura, M.
TITLE PRODUCTION OF HUMAN APOLIPOPROTEIN E-LIKE PROTEIN
JOURNAL Patent: JP 1986096997-A 1 15-MAY-1986;
MITSUBISHI CHEM IND LTD
COMMENT
OS Human (Homo sapiens)
PN JP 1986096997-A/1
PD 15-MAY-1986
PF 16-OCT-1984 JP 1984216987
PI TERANISHI YUTAKA, MATSUI YASUSHI, IKEDA YASUOKO, KIMURA MASAKO
PC C12P21/00,A61K37/4,A61K37/04,C12N15/00,(C12P21/00,C12R1:19),
PC (C12N15/00,
PC C12R1:19);
CC strandedness: Double;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: tissue_type=Liver;
key location/Qualifiers
FT 3'UTR 1..14
FT sig_peptide 15..68
FT FT /product='human apolipoprotein E signal' FT

FEATURES
source 1..1110
location/Qualifiers
BASE COUNT 198 a 353 c 416 g 143 t
ORIGIN

Alignment Scores:
Pred. No.: 1,22e-82 Length: 1110
Score: 1577.00 Matches: 316
Percent Similarity: 99.68% Conservative: 0
Best Local Similarity: 99.68% Mismatches: 1
Query Match: 99.24% Indels: 0
DB: Gaps: 0

US-09-827-854-16 (1-317) x E00823 (1-1110)
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 15 ATGAAGGTTCTGGGCTCGTTCGTGCTGACATTCCTGGCAGGATGCCAGGCCAAGGTG 74
QY 21 GUGGAlaValAlaGlnThrGluProGluLeuArgGlnGlnThrGlnThrGlnThr 40
DB 75 GAGCAAGCGGTGGAGACAGCCGGAGCCGAGCTGCCAGCAGACCGAGTGGCAGACC 134
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 135 GGCACAGCGCTGGGAACCTGGCAGCTGGCTCTTTGGGATTACCTGGCGTGGCAGACA 194

```


QY	61	LeuSerGIuGIuValGIuGIuLeuLeuSerSerGIuValThcGIuLeuAArgAla	80
Db	195	CGTCTGACAGAGGTGCAGAGAGAGCTGCTACACTCCAGTACACCGAGACTGAGGGG	254
QY	81	LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGIuLeu	100
Db	255	CGATGTGACAGAACCCATGTGAAGAGTGTGAAGGGCTCAAAATCGCACTGAGAGACAACTG	314
QY	101	ThrProValAlaGIuGIuThrThrArgAlaArgLeuSerLysGIuLeuGIuAlaGIuAla	120
Db	315	ACCCCGGTGGCGGAGAGACCGCGGCGAGCGCTGTCCAAAGAGACTGACGCGCGCACGGCC	374
QY	121	ArgLeuGIuAlaAspMetGIuAspValCysGIuArgLeuValGIuTyrArgGIuVal	140
Db	375	CGGCTGTGGCGGACATGTGAGAGACGTGTGTGGCCGCGCTGTGCATGACCGGGCGAGTGTG	434
QY	141	GIuAlaMetLeuGIuLysInSerThrGIuLeuLeuArgValArgLeuAlaSerHisLeuArg	160
Db	435	CAGGCGATCTCTGGCGCAGACACCGAGGAGCTCGCGGTCCGCTCCACCTCGGCC	494
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGIuLysCysLeuValAlaTyr	180
Db	495	AAGCTGCTGAAAGCGGCTCTCCGCGATGCCATGACCTCAAGAGCGCTGCGACAGTAC	554
QY	181	GIuAlaGIuAlaArgGIuGIuValGIuGIuValGIuLeuSerAlaIleArgGIuArgLeuGIu	200
Db	555	CAGGCGCGGGCGCGAGAGGCGCGCGAGCGCGGCTCAGCGCATCCGAGCGGCGTGGGG	614
QY	201	ProLeuValGIuGIuGIuGIuArgValArgAlaAlaThrValGIuLysLeuAlaGIuPro	220
Db	615	CCCCTGTGTGAACAAGGCGCGTGTGGCGCGGCTGCTGTGCTCCCTGGCGCGGACGCG	674
QY	221	LeuGIuGIuArgAlaGIuAlaTyrPGLysGIuArgLeuArgAlaArgMetGIuLysMetGIu	240
Db	675	CTACAGAGAGCGGGGCCAGAGGCTGTGGGGCGAGCGCGCTGCGCGCGGATGAGAGATGGCG	734
QY	241	SerArgThrArgAspArgLeuAspGIuValLysGIuGIuAlaGIuAlaGIuValArgAlaLys	260
Db	735	ACCCGAGACCCCGAGCGCGCTGTGACAGAGTGAAGAGAGAGAGTGGCGAGTGGCGGCCCAAG	794
QY	261	LeuGIuGIuGIuAlaGIuGIuGIuLysLeuArgLeuGIuAlaGIuAlaPheGIuAlaArgLeuLys	280
Db	795	CTGGAGAGAGAGAGCCAGACAGATACGCTGTGACGCGCGAGGCTTCCAGGCGCGGCTTCAAG	854
QY	281	SerTyrPheGIuProLeuValGIuAspMetGIuArgIleTyrPAlaGIuLysValGIuLys	300
Db	855	ACCTGATTCGAGACCCCTGTGTGGAACAATGACAGCGCCAGTGGGCGGCGTGTGAGAAG	914
QY	301	ValGIuAlaAlaValGIuThrSerLysAlaIleProValProSerAspAsnHis	317
Db	915	GTGACAGGCTCGGTGGGACACAGCGCGCTGTGTGCCAGCGACAATCAC	965
RESULT 4			
LOCUS	AX302545	1147 bp	DNA
DEFINITION	Sequence 63 from Patent W00175177.		
ACCESSION	AX302545		linear
VERSION	AX302545.1		
KEYWORDS	GI:17383082		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1		
AUTHORS	Motlin, P.J., Sherman-Baust, C.A., Pizer, E.S. and Hough, C.D.		
TITLE	Tumor markers in ovarian cancer		
JOURNAL	Patent: WO 0175177-A 63 11-OCT-2001.		
FEATURES	THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)		
	Location/Qualifiers		
	1..1147		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		

BASE COUNT	210 a	365 c	425 g	147 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	1,26e-82	Length:	1147		
Score:	1577.00	Matches:	316		
Percent Similarity:	99.68%	Conservative:	0		
Best Local Similarity:	99.68%	Mismatches:	1		
Query Match:	99.24%	Indels:	0		
DB:	6	Gaps:	0		
US-09-827-854-16 (1-317) x AX302545 (1-1147)					
QY	1	MethylsValleuITrPalalalLeuLeuValThrheLeuAlaGlyCysGlnAlaIalysVal	20		
DB	46	ATGAAGGTTTCGTGGGCTGCTGTCTGTCTCAATTCCTGGCAGGATCCAGGCCAAAGTG	105		
QY	21	GIUGlnAlaValAGluThrGluProGluProGluLeuAarGlnGlnThrGluITrPGlnSer	40		
DB	106	GAGCAAGCGGCTGGAGACAGACACCCGAGCTCGGCCACAGACAGACGATGGCAGAC	155		
QY	41	glyGlnArGTrPGluLeuAlaLeuGlyArGheTrpAspTrpLeuArGTrpValGlnThr	60		
DB	166	GGCCAGGCGCTGGAACTGGCACTGGGTGCGTTTGGATTACCTGGCGGTGGTGGACAGA	225		
QY	61	LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArGala	80		
DB	226	CTGTCTGGAGCGTGGCAGAGAGACTCTCAAGCTCCAGGTCACCCAGGAACGTAGAGCG	285		
QY	81	LeuMeTaSPgluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGluGlnLeu	100		
DB	286	CTGATGGAGCGAGACCATGAAAGAGTTGTAAGGCCCTTCAAAATCGGAACCTGGAGAACTG	345		
QY	101	ThrProValAlaGluGluGluThrArGalaArGLeuSerLysGluLeuGlnAlaAlaGlnAla	120		
DB	346	ACCCGCGTGGCGGAGAGAACCCGCGGCACGCGCTGTCCAAGAGAGCTGCAGCGCGCAGGCC	405		
QY	121	ArgLeuGlyAlaAspMetGluAspValCysGlyArGLeuValGlnTrpArGlyGluVal	140		
DB	406	CGGCTGGGCGCGAGATGAGAGAGCTGGCGGCCCTGTGTGACGTACCCGCGGAGGTG	465		
QY	141	GlnAlaMeTleuGlnGlnSerThGluGluLeuAarGValArGLeuAlaSerHisLeuArG	160		
DB	466	CAGGCAATGCTGGGCGAGACACCGAGAGAGCTGGGGTGGCGCTGCCACACTGGCG	525		
QY	161	LysLeuArGlySarGLeuLeuArGAspAlaAspAspLeuGlnLysCysLeuAlaValTrp	180		
DB	526	AAGTGCCTAAGCGGCTCTCTCCGAGATGCCATGACCTTGCAGAAAGCCCTGGCAGTGTAC	585		
QY	181	GlnAlaGlyAlaArGgGluGluGlyAlaGluAarGlyLeuSerAlaIleArGgLuArGLeuGly	200		
DB	566	CAGGCGGGGGCCCGGAGAGGGCGCGAGAGCGGCTTCAGCCGCAATCCCGAGCGCTGGGG	645		
QY	201	ProLeuValGluGlnGlnGlyArGValArGAlaAlaIleArGlySerLeuAlaGlnPro	220		
DB	646	CCCCGTGGGAACAAGGGCGCGGTGGGGCCGCCACTGTGGGGTCTCCGTGGCGGCAAGCG	705		
QY	221	LeuGlnGlnArGAlaGlnAlaITrPGlyLuArGLeuArGAlaArGmetGluGluMetGly	240		
DB	706	CTACAGGAGCGCGGCCACAGGCTCTGGGGCGAGGCGCTGGCGCGCGGAGTGGAGATGGCG	765		
QY	241	SerArGTrArGAspArGAspArGLeuAspGluValLysGlnGlnValAlaGluValArGAlaLys	260		
DB	766	AGCCGAGCCCGCGACCGCTTGGAGAGGTGAAGAGCAAGGTGGCGAGGTGCCGCCAAG	825		
QY	261	LeuGlnGlnGlnAlaGlnGlnGlnIleArGLeuGlnAlaGlnAlaIlePheGlnAlaArGLeuLys	280		
DB	826	CTGGAGGAGCGAGGCCCAAGACATACGCTCGAGGCGCAGAGCCCTTCCAGGCCCGCCCTCAAG	885		
QY	281	SerTrpPheGluProLeuValGluAspMetGlnArGLeuITrPalalalGlyLeuValGluLys	300		
DB	886	AGCTGGTTCGAGCCCTGGTGGAGAACATCAGCGCCACAGTGGCGCGGCTGGTGGAGAAAG	945		

Oy 301 ValGlnAlaIaIaValGlyThrSerAlaIaIaProValProSerAspAsnHis 317
 |||||
 Db 946 GTGACAGGCTGGCGTGGGACACAGCGCCGCTGTGCCAGAGCAATCAC 996
 |||||
 RESULT 5
 BD004278 1156 bp DNA linear PAT 31-JAN-2002
 LOCUS Apo E humanized mammal
 DEFINITION BD004278
 ACCESSION BD004278.1 GI:18632239
 VERSION JP 2001017028-A/2.
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens.
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1156)
 Fujita,S., Hamanaka,H.,
 Apo E humanized mammal
 Patent: JP 2001017028-A 2 23-JAN-2001;
 JOURNAL MITSUBISHI CHEMICAL CORP
 OS Homo sapiens (human)
 PN JP 2001017028-A/2
 PD 23-JAN-2001
 PF 28-APR-2000 JP 2000128919
 PR
 PI SHINOBU FUJITA,HIROKI HAMANAKA,YUKO FUKUI,MINESUKE YOKOYAMA PC
 A01667/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10, PC
 C12N15/09//C07K14/775,
 PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,(C12N5/00,C12R1:91) CC
 FEATURES
 source location/Qualifiers
 FT Key CDS location/Qualifiers
 (61)..(1011).
 1..1156
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 208 a 368 c 432 g 148 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.27e-82 Length: 1156
 Score: 1577.00 Matches: 316
 Percent Similarity: 99.68% Conservative: 0
 Best Local Similarity: 99.68% Mismatches: 1
 Query Match: 99.24% Indels: 0
 DB: Gaps: 0
 US-09-827-854-16 (1-317) x BD004278 (1-1156)
 Oy 1 MetLysValLeuTrpAlaIaIaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||||
 Db 61 ATGAAGGTTCTGTGGGCTGGCTGCTGTGCATTCCTGGCAGGTGCCAGGCCAAGGTG 120
 |||||
 Oy 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 |||||
 Db 121 GAGCAAGCGGTGAGACAGAGCGGAGCCGAGCTGCGCCAGCAACGAGTGGCAGAGC 180
 |||||
 Oy 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 |||||
 Db 181 GGCCAGCCTGGGAACCTGGCACTGGTGTGCTTGGGATTAACCTGCGTGGTGCAGACA 240
 |||||
 Oy 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||||
 Db 241 CTGTGTGACGAGGTGACAGAGAGCTGCTCACTCCAGGTCCAGGACCAAGTGAAGCGC 300
 |||||
 Oy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnGluLeu 100
 |||||
 Db 301 CTGATGACGAGACCATGAGAGAGTTGAAGGCTTCAATGGAATCGAGAGCAACTG 360
 |||||
 Oy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaIa 120
 |||||
 Db 361 ACCCGGTGGCGGAGAGACGCGGCGCTGTCCAAAGGAGAGTCCAGGCGCGCAGGCC 420
 |||||

Oy 121 ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnTrpArgGlyVal 140
 |||||
 Db 421 CGGCTGGGCGCGAGATGAGAGACGTGTGGCCGCTGTGTACATACCGGCGAGGTG 480
 |||||
 Oy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||
 Db 481 CAGGCCATGCTCGGCCACAGACACCGAGAGAGCTGGGGTGGCCCTGCCACCTGGCGC 540
 |||||
 Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 |||||
 Db 541 AAGCTGCTTAAGCGGCTCTCCGCGATGCGATGACCTGCACAAAGCGCTGTGCACTGTAC 600
 |||||
 Oy 181 GlnAlaGlyAlaArgGluGlyValGlnArgGlyLeuSerAlaIleArgGluTrpGluGly 200
 |||||
 Db 601 CAGGCGGGGCGCGGAGAGGCGCGAGCGGCTCAGCGCCATCCGAGAGGCGCTGGG 660
 |||||
 Oy 201 ProLeuValGluGlnGlyArgValArgAlaAlaIleThrValGlySerLeuAlaGlnPro 220
 |||||
 Db 661 CCCCTGTGGAAACAGCGCGCGCTGGCGCGCCACTGTGGCTCCCTGGCGCCAGCGC 720
 |||||
 Oy 221 LeuGlnGluArgAlaGlnAlaIleTrpGlyGluArgLeuArgAlaArgMetGluGluTyr 240
 |||||
 Db 721 CTACAGAGAGCGGCGCCAGGCTGGGGCGAGCGGCTGGCGCGGATGGAGAGATGGGC 780
 |||||
 Oy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 |||||
 Db 781 AGCGGAGACCGCGGACCGCTGTGACAGAGTGAAGAGACAGTGGCGAGGTGGCCCAAG 840
 |||||
 Oy 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 |||||
 Db 841 CTGGAGAGAGAGCGCCACAGATACGCTGCAGCGCGAGGCTTCAGAGCGCGCTCAAG 900
 |||||
 Oy 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValLys 300
 |||||
 Db 901 AGCTGTGTGACCGCTGTGGTGAAGACATGACGCCCATGTGGCGGTGTGTGAGAG 960
 |||||
 Oy 301 ValGlnAlaIaIaValGlyThrSerAlaIaIaProValProSerAspAsnHis 317
 |||||
 Db 961 GTGACAGGCTGGCGTGGGACACAGCGCCGCTGTGCCAGAGCAATCAC 1011
 |||||
 RESULT 6
 HUMAPOE3 1156 bp mRNA linear PRI 24-NOV-2000
 LOCUS Homo sapiens preapolioprotein E (APOE) mRNA, complete cds.
 DEFINITION K00396
 ACCESSION K00396.1 GI:178850
 VERSION
 KEYWORDS apolioprotein; apolioprotein E; lipoprotein; polymorphism; very low density lipoprotein.
 SOURCE Homo sapiens.
 Homo sapiens.
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 355 to 1156)
 Breslow,J.L., McPherson,J., Nussbaum,A.L., Williams,H.W.,
 Iofquist-Kahl,F., Karathanasis,S.K. and Zannis-V.I.
 Identification and DNA sequence of a human apolioprotein E cDNA
 clone
 JOURNAL J. Biol. Chem. 257 (24), 14639-14641 (1982)
 MEDLINE 83082756
 PUBMED 6897404
 2 (bases 250 to 777)
 Wallis,S.C., Rogne,S., Gill,L., Markham,A., Edge,M., Woods,D.,
 Williamson,R. and Humphries,S.
 The isolation of cDNA clones for human apolioprotein E and the
 detection of apoE RNA in hepatic and extra-hepatic tissues
 JOURNAL EMBO J. 2 (12), 2369-2373 (1983)
 MEDLINE 84131952
 PUBMED 6199196
 3 (bases 1 to 1156)
 Zannis,V.I., McPherson,J., Goldberger,G., Karathanasis,S.K. and
 Breslow,J.L.
 Synthesis, intracellular processing, and signal peptide of human
 apolioprotein E

JOURNAL J. Biol. Chem. 259 (9), 5495-5499 (1984)
 MEDLINE 84185684
 PUBMED 6325438
 REFERENCE 4 (bases 88 to 1156)
 AUTHORS McLean,J.W., Elshourbagy,N.A., Chang,D.J., Mahley,R.W. and Taylor,J.M.
 TITLE Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing of a new variant
 JOURNAL J. Biol. Chem. 259 (10), 6498-6504 (1984)
 MEDLINE 84212473
 PUBMED 6327682
 REFERENCE 5 (bases 577 to 624)
 AUTHORS Gill,L.U., Peoples,O.P., Pearson,D.H., Robertson,F.W., Humphries,S.E., Cumming,A.M. and Hardman,N.
 TITLE Isolation and characterisation of a variant allele of the gene for human apolipoprotein E
 JOURNAL Biochem. Biophys. Res. Commun. 130 (3), 1261-1266 (1985)
 MEDLINE 85279526
 PUBMED 2992507
 REFERENCE 6 (sites)
 AUTHORS Rall,S.C. Jr., Newhouse,Y.M., Clarke,H.R., Weisgraber,K.H., McCarthy,B.J., Mahley,R.W. and Bersot,T.P.
 TITLE Type III hyperlipoproteinemia associated with apolipoprotein E phenotype E3/3. Structure and genetics of an apolipoprotein E3 variant
 JOURNAL J. Clin. Invest. 83 (4), 1095-1101 (1989)
 MEDLINE 89198059
 PUBMED 2539388
 COMMENT [1] corrected in [J. Biol. Chem. 258, 11422-11422 (1983)]. [J. Biol. Chem. 258, 11422-11422 (1983)] correction of [1]. [4] epsilon-3 and variant.
 [5] epsilon-2 allele.
 [6] sites; mutations resulting in type III hyperlipoproteinemia. Apo E is a component of normal human very low density lipoprotein. There are six human apo E phenotypes known to result from a single structural gene, three of the common alleles being epsilon-4, epsilon-3 and epsilon-2. This sequence appears to be of the epsilon-3 allele. [1] argues that the apo E polymorphism involves mutations in the structural coding region; for example the epsilon-2 phenotype which is characterized by hyperlipoproteinemia is thought to result from a c to t change (arg to cys) at base 586 below [3]. [5]. The sequence shown is 578 homologous with human apo A-I and 818 homologous with rat apo E. For the epsilon-4 sequence, see the separate entry.
 [J. Biol. Chem. 258, 11422-11422 (1983)] is too extensively revised relative to [1] to record the
 revisions in the FEATURES table. The sequence below is that of [J. Biol. Chem. 258, 11422-11422 (1983)]

Apo E is located on chromosome 19 --Jackson, Bruns and Breslow, PNAS USA 81, 2945-2949 (1984) -- and is believed to be linked to the Apo C-II gene (see separate entry).
 The two mutations causing type III hyperlipoproteinemia (apolipoprotein E phenotype E3/3) produces substitutions of Arg for Cys at amino acid 112 and Cys for Arg at amino acid position 142. Draft entry and printed copy of sequence for [1] kindly provided by L.L. Gill, University of Aberdeen.
 Complete source information:
 Human liver [1], [J. Biol. Chem. 258, 11422-11422 (1983)], [2], [3], [4] and blood [5], cDNA to mRNA.

FEATURES

source location/Qualifiers

1..1156

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="19q13.2"

/tissue.type="liver and blood"

1..1156

/gene="APOE"

61..1014

/gene="APOE"

/codon.start=1

/product="preapolipoprotein E"

/protein_id="AAB59546.1"

sig_peptide

mat_peptide

variation

variation

variation

variation

variation

variation

variation

variation

variation

variation

variation

variation

/db_xref="GI:178851"
 /db_xref="GDB:G00-119-691"
 /translation="MKVLAALLTFFLAGCAQKVEQAVETPEPELROOTEMQSGORM
 ELAIGRFMDYLVQVOTLSROVQEEISSQVTOELRAIMDETKEELKAYKSELEGLTP
 VAEFRRARSKELQAQAARICADMDYCGRIYQVQAGVQAMGOSTEELRVRLASHLR
 KIKRRLRDADLQRLAVYQAGARGAERGLSATIRERGLPYEGRRAATVGSFLAG
 QPLQERQAQMGERRLARMEGSRTRDRIDVEKEQVAEVRALLQEQAOQIRLQABAFQ
 ARLKSWFEPLVEDMQRQVAGLVEKVOAAVGTSAAPVPSDNH"
 61..114

/gene="APOE"
 /note="apolipoprotein E"

115..1011

/gene="APOE"

/product="apolipoprotein E mature peptide"

267

/gene="APOE"

/note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3], [4]"

/citation=[2]

/replace="c"

279

/gene="APOE"

/note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3]"

/citation=[4]

/replace="a"

288

/gene="APOE"

/note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3]"

/citation=[4]

/replace="a"

369

/gene="APOE"

/note="in variant"

/citation=[4]

/replace="a"

409

/gene="APOE"

/note="in variant"

/citation=[4]

/replace="a"

448

/gene="APOE"

/note="in type III hyperlipoproteinemia"

/replace="c"

538

/gene="APOE"

/note="in type III hyperlipoproteinemia"

/replace="t"

568

/gene="APOE"

/note="in variant"

/citation=[4]

/replace="c"

586

/gene="APOE"

/note="in epsilon-2 allele"

/citation=[3]

/citation=[5]

/replace="t"

702

/gene="APOE"

/note="c in [J. Biol. Chem. 258, 11422-11422 (1983)], [3], [4]"

/citation=[2]

/replace="g"

783

/gene="APOE"

/note="in variant"

/citation=[4]

/replace="t"

858

/gene="APOE"

/note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3]"


```

and [4] variant: a in [4] 'normal'
/citation=[4]
/replace="a"

BASE COUNT      208 a      368 c      432 g      148 t
ORIGIN

Alignment Scores:
Pred. No.:      1,27e-82      Length:      1156
Score:          1577.00      Matches:      316
Percent Similarity: 99.68%      Conservative: 0
Best Local Similarity: 99.68%      Mismatches: 1
Query Match:      99.24%      Indels:      0
DB:               Gaps:      0

US-09-827-854-16 (1-317) x HUMAPOE3 (1-1156)

OY      1  MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
      |||||||
DB      61  ATGAAGGTTCTGTGGCGCTGCTGTGTCACATTCTGGCAGATGCGCAGGCCAAGGTG 120
      |||||||

OY      21  GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGluThrGluThrPglInser 40
      |||||||
DB      121  GAGCAACGCGTGGAGACAGACGCCGAGCCGAGCTGCCAGACAGACGAGTGGCAGAGC 180
      |||||||

OY      41  GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
      |||||||
DB      181  GGCCAGCGCTGGGAACCTGGCACTGGGTGCTTTTGGGATTAAGTGGTGGTGGTGGTGGTGG 240
      |||||||

OY      61  LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
      |||||||
DB      241  CTGTGTGAGCAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
      |||||||

OY      81  LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGlnLeuGlnGluGlnGln 100
      |||||||
DB      301  CTGATGAGCAGACCATGATGAGAGAGTGTGAGGCTTACAAATCGGAAGTGGAGACAACTG 360
      |||||||

OY      101  ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
      |||||||
DB      361  ACCCGGCGGGGAGAGAGAGCGGGGACGCGCTGCCAAGAGACTGCGAGCGCGCGCAGCGC 420
      |||||||

OY      121  ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
      |||||||
DB      421  CGCGTGGGCGCGAGCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
      |||||||

OY      141  GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerLysLeuArg 160
      |||||||
DB      481  CAGGCGCATGCTGGCCAGAGACCGCAGAGAGCGCGGCTGCGCTGCCCTCCACCTGGCGC 540
      |||||||

OY      161  LysLeuArgGlyArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
      |||||||
DB      541  AAGCTGGCGTAAGCGGCTCTCCGCGATGCCGATGACCTGCAGAACCGCGCTGGCAGTGTAC 600
      |||||||

OY      181  GlnAlaGlyAlaArgGluGluGluGluAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
      |||||||
DB      601  CAGGCGCGGGCGCGCGAGAGCGCGCGAGCGCGCTCAGCGCATCCGCGAGCGCGCTGGGG 660
      |||||||

OY      201  ProLeuValGluGlnGlyArgValArgValArgAlaThrValGlySerLeuAlaGlyLysPro 220
      |||||||
DB      661  CCCCTGGTGGAGACAGAGCGCGCGCTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
      |||||||

OY      221  LeuGlnGluArgAlaGlnAlaAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
      |||||||
DB      721  CTACAGAGAGCGGGCGCGCGAGCGCGCGCTGCGCGCGCGCGATGGAGAGAGATGGCGC 780
      |||||||

OY      241  SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
      |||||||
DB      781  AGCCCGAGCGCGCGCGCGCTGGAGCAGAGTGAAGGAGAGCGAGTGGCGAGTGGCGCGCAG 840
      |||||||

OY      261  LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
      |||||||
DB      841  CTGGAGAGAGAGCGCGCGCGAGCATACGCTTGGAGCGCGAGGCTTCCAGGCGCGCTCAAG 900
      |||||||

OY      281  SerTrpPheGluProLeuValGluLysPheGlnArgGlnTrpAlaGlyLeuValGluLys 300
      |||||||

```

```

|||||
DB      901  ACCTGCTCGACCCCTGCTGAGACATGACAGCCCGAGTGGCGCGCGCTGCTGAGAGAG 960
      |||||||

OY      301  ValGlnAlaValAlaGlyThrSerAlaIleProValProSerAspAsnHis 317
      |||||||
DB      961  GTTCAGAGCTGCCGTGGGACACAGCGCGCGCTGTGTGCCAGCAATCAC 1011
      |||||||

```

```

RESULT 7
BC003557      1186 bp      mRNA      linear      PRI 12-JUL-2001
LOCUS      BC003557
DEFINITION      Homo sapiens, apolipoprotein E, clone MGC:1571 IMAGE:3355712, mRNA,
complete cds.
ACCESSION      BC003557.1  GI:13097698
VERSION      MGC.
KEYWORDS
SOURCE
ORGANISM      Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1186)
REFERENCE
AUTHORS      Strausberg, R.
TITLE      Direct Submission
JOURNAL
Submitted (20-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

```

```

REMARK
COMMENT
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
BC DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
Info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Beta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeay, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schelin, Diane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasha van den Bosch, Jill Vardy,
George Yang, Scott Zuyerdun, Marco Marra.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 6 Row: h Column: 23
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4557324.

```

FEATURES
SOURCE
1..1186
/organism="Homo sapiens"
/db_xref="locusid:348"
/db_xref="taxon:9606"
/clone="MGC:1571 IMAGE:3355712"
/tissue.type="Eye, retinoblastoma"
/clone.lib="NIH_MGC_16"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
46..999
/codon.start=1
/product="apolipoprotein E"
/protein_id="AAH03557.1"
/db_xref="GI:13097698"
/translation="MKVLMALVTFLPAGCOAKVEQAVETPEPELNLQRTWESGGRN
ELAGRWDTLRWVOTLSEVOBELLSOVQBELRALMDETMKELAKYKSLLEQLTP
VAETTRARLSKEILOAAOARLQADMEDYCGRLVQRYGEVOALQSTELRRLASLSLR
KLRRLRDADDLQRLAVVOAGAREGERSLAIREFGLVQGRVRAATVSLG
OPLQEROAWEERLRARMEGSRTRIDRDKROVAEVRKLEQAOQIRLQAEARO
ARKLSWEPLVEDQROMAGCIYVEVQAAVGSAPVPSDNH"

```

```

BASE COUNT      248 a      366 c      425 g      147 t
ORIGIN

```


Db 315 ACCCGGTGGCCGAGAGACGCCGGCCAGCGCTGTCCACAGAGCTGCAGGCGGCCAGGCC 374
OY 121 ArgLeuGluYAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 375 CGGCTGGGGCGGACATGGAGGACGTGTGCGGCGCGCTGTGTGACAGTACCGCGCCAGAGTG 434
OY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 435 CAGGCGCATCTCGGCGCAGAGCACCGGAGAGTGGGTCGCGCTGCGCTCCACCTGCCG 494
OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 495 AAGCTCGGTAAACGGCTCTCCCGCATGGCCCATGACTCTGCAAGAGCGCTGGCAGGTATC 554
OY 181 GlnAlaGlyAlaArgGluGluValArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 555 CAGGCGGGGGCGCCGAGGGCGCCGAGCGCGCTCAGCGCATCCGCGAGCGCTGGGG 614
OY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 615 CCCGTGTGGAAACAGGCGCGGTGGGGCGCCAGCTGTGGCTCCCTGGCGGCCAGCGC 674
OY 221 LeuGlnGluArgAlaGlnAlaThrPgluArgLeuArgAlaArgMetGluGlnMetGly 240
Db 675 CTACAGGAGCGGGCGCCAGGCTGGGGCGAGCGCGCTGGCGCGCATGGAGAGATGGGC 734
OY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 735 ACCCGGACCGCGGACCGCTGGACGAGGTGAGAGAGAGGTGGCGAGGTGGCGCCAAAG 794
OY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db 795 CTGGAGAGACAGCCGACAGATACGCTGCAGGCGGAGGCTTCAGGCGCGCGCTCAAG 854
OY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 855 AACTGTGTGAGCCCTGTGTGAGAGCATGACAGCCGAGGTGGCGCGCTGTGTGAGAG 914
OY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 915 GTGCAGGCTGCCGTGGGACACGCGCGCCCTGTGTGCCAGGACATATC 965
RESULT 9
AX333278 1157 bp DNA linear PAT 09-JAN-2002
LOCUS AX333278
DEFINITION Sequence 3787 from Patent WO0194629.
ACCESSION AX333278
VERSION AX333278.1 GI:18123912
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Young, P.E., Augustus, M., Carter, R.C., Ebnner, R., Endress, G.,
Horrikan, S., Soppet, D.R., and Weaver, Z.
Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 3787 13-DEC-2001;
FEATURES
source
1. 1157
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN

Alignment Scores:

Pred. No.: 4.2e-82
Score: 1568.00
Percent Similarity: 99.05%
Best Local Similarity: 99.05%
Query Match: 98.68%

Length: 1157
Matches: 314
Conservative: 0
Mismatch: 3
Indels: 0

DB: 6 Gaps: 0
US-09-827-854-16 (1-317) x AX333278 (1-1157)
OY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAGGTTCTGTGGGCTGGCTGTGGTGTGATTCCTGTGGAGATGCCAGGCCAAGGTG 121
OY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 122 GAGCAAGCGGTGGAGCAAGCGGAGGCCGAGCTGCGCAGACAGACAGAGTGGCAGAGC 181
OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 182 GGCCAGCGCTGGAGACTGGCATTGGGTGCTTTGGGATTACCTGCGCTGGGTGCAGACA 241
OY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 242 CTGTCTGAGCAGGTGACAGAGAGAGCTGCTCAGCTCCCAAGTCACCCAAAGACTGAGGCGC 301
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
Db 302 CTGATGGACAGACCATGAAAGAGTTGAGGCTTCAAAATTCGAACTGGAGAGAACTTG 361
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 362 ACCCGGTAGCGGAGAGAGCGGGCAGCGCTGTCCAAAGAGCTGCAGAGCGCGCAGGCGC 421
OY 121 ArgLeuGluYAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 422 CGGCTGGGGCGGACATGAGAGAGCTGTGGCGCCCTGTGTGATACCGCGCGCAGAGTG 481
OY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCGCATCTGGCGCAGAGCACCGAGAGCTGGGGGTGGCTGTGCTCCACCTGGCGC 541
OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 542 AAGCTGCGTAAACGGCTCTCCCGCATCCGATGACCTGCAGAACCGCTGCGCATGTATC 601
OY 181 GlnAlaGlyAlaArgGluGluValArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 602 CAGGCGGGGGCGCGGAGGCGCGCAGCGCGCTCAGCGCATCCGCGAGCGCGCTGGGG 661
OY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 662 CCCGTGTGGAAACAGGCGCGGTGGGGCGCCAGCTGTGGCTCCCTGGCGGCCAGCGC 721
OY 221 LeuGlnGluArgAlaGlnAlaThrPgluArgLeuArgAlaArgMetGluGlnMetGly 240
Db 722 CTACAGAGAGCGGGCGCCAGGCGCTGGGGCGAGCGGCTGCCGCGCATGGAGAGATGGGC 781
OY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 782 AGTCGAGACCGCGCAGCGCTGTGAGAGGTGAAAGGAGAGGTGGCGAGAGTGGCGCCCAAG 841
OY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db 842 CTGGAGAGAGCGGCCAGACAGTACGCTGCAGGCGCGAGGCTTCAGGCGCGCTCAAG 901
OY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 902 AAGCTGTTCAGAGCCCTGTGTGAGAACATGACAGCGCCAGTGGCGCGCTGTGTGAGAG 961
OY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGCAGGCTGCGTGGGACACGAGCGCGCCCTGTGTGCCAGGACATATC 1012
RESULT 10
AX409597 1157 bp DNA linear PAT 14-JUN-2002
LOCUS AX409597
DEFINITION Sequence 2244 from Patent WO0229103.
ACCESSION AX409597

VERSION	AX0409597.1	GI:214442302
KEYWORDS		
SOURCE	human.	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE		
AUTHORS	1 Alvaras, C., Horne, D., Pears-da-Silva, S. and Vockley, J.G.	
TITLE	Gene expression profiles in liver cancer	
JOURNAL	Patent: WO 0229103-A 2244 11-Apr-2002;	
FEATURES	GENE LOGIC INC (US)	
source	location/Qualifiers	
	1..1157	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/note="EMBL/GenBank Accession No. M12529"	
BASE COUNT	212 a 370 c 426 g 149 t	
ORIGIN		
Alignment Scores:		
Pred. No.:	4.2e-82	1157
Score:	1568.00	Matches: 314
Percent Similarity:	99.05%	Conservative: 0
Best Local Similarity:	99.05%	Mismatches: 3
Query Match:	98.68%	Indels: 0
DB:	6	gaps: 0
US-09-827-854-16 (1-317) x AX0409597 (1-1157)		
QY	1 MettysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal	20
DB	62 ATGAAGGTTCTGTGGGCTGCGTGTGTCATTCCTCGCAGAGATGGCAGCCAAAGTG	121
QY	21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrPglIns	40
DB	122 GAGCAAGGGGTGGAGACAGACCAGGAGCCAGCTGGCCAGCAGAGAGAGTGGCAGAGC	181
QY	41 GlyGlnAqTTPgluLeuAlaLeuGlyYArgPheTrpAspTyrLeuArgTTPValGlnThr	60
DB	182 GGCCAGCGCTGGGAAGTGGCACTGGGTGCTTTTGGATTACCTGCGCTGGGTGCAGACA	241
QY	61 LeuSerGlnGlnValGlnGluGluLeuLeuSerGlnValThrGlnGluLeuArgAla	80
DB	242 CTGTCTGACGAGTGTACAGAGAGACTGCTGCTCAGTCCCAAGTCAACCCAAAGAACTGAGGGCG	301
QY	81 LeuMetAspGluThrMetIysGluLeuIysAlaIleTyrIysSerGluLeuGluGlnLeu	100
DB	302 CTGATGGACGAGACCATTAAGAGACTTGAAGGCTTCAAAATCGGAATCGGAGAACAACTG	361
QY	101 ThrProValAlaGluGluThrArgAlaArgLeuSerIysGluLeuGlnAlaAlaGlnAla	120
DB	362 ACCCGGTAGCGGAGAGAACACCGGGGACAGGCTGTCCAAAGAGCTGCAGAGCGCCAGGCC	421
QY	121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal	140
DB	422 CGGGTGGCGCGGACATGAGAGAGCTGGCGCGCCCTGTGGTACCGCGGCGAGGTG	481
QY	141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg	160
DB	482 CAGGCAATGCTCGGCGCAAGACACCGAGAGAGTGGGGGGCGCCCTCCACCTTCGCGC	541
QY	161 LysLeuArgIysArgLeuLeuArgAspAlaAspAspLeuGlnIysCysLeuAlaValTyr	180
DB	542 AAGCTGCTAAGCGGCTCTCTCGGAGTCCGATACCTCGCAAGAGCCCTGGCAGTGTAC	601
QY	181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200
DB	602 CAGGCGGGGGCGCGAGGGGCGCGAGCGGGCGCTTCAGCGCATCCCGAGCGCTGGGG	661
QY	201 ProLeuValGluGlnGlyArgValArgAlaAlaIleThrValGlySerLeuAlaGlyGlnPro	220
DB	662 CCCCTGTGGTGAACAGGGCGCGTCCGGGCGCCACTGTGGGTCTCCCTGGCGGGCAGCCG	721

QY	221	leunglmglaargalaglalnlaatrprgljglaargleuargalaaargmgluimwety	240
QY	221	leunglmglaargalaglalnlaatrprgljglaargleuargalaaargmgluimwety	240
Db	722	CTACAGGAGAGGGGGCCAGGCGCTGGGGGCGAGCGGCTGGCGCCGGATGGAGGATGGGC	781
QY	241	SerATgThrArGAsPArgLeuAsPGLuValLysGluGlnValAlaGluValAlaGluLys	260
Db	782	ACTCGGACCCCGGACCGGCTGGACGAGTGAAGAGACAGAGTGGCGGAGTGGCGGCCAAG	841
QY	261	leuglmglaargalaglalnlaatrprgljglaargleuargalaaargmgluimwety	280
Db	842	CTGAG	901
QY	281	SerTPrpHeGluProLeuValGluLysPheGlnAlaArgLeuValAlaGluLys	300
Db	902	ACCTGGTTCGAGCCCGCTGGTGGAGACATGACAGCCCGCATGGGGCGGGCTGGTGGAGAG	961
QY	301	ValGlnAlaAlaValAlaGluThySerAlaAlaProSerAspAsnHis	317
Db	962	GTGCAAGGCTGGCTGGGACACAGCGCCGCGCTGTGGCCAGGACAAATCAC	1012
RESULT	11		
LOCUS	115975	1157 bp	DNA linear PAT 03-APR-1996
DEFINITION	Sequence 1 from patent US 5472858.		
ACCESSION	115975		
VERSION	115975.1	GI:1250883	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1157)		
AUTHORS	Atlie,A.D., Gretch,D.G., Sturley,S.L. and Beckage,N.E.		
TITLE	Production of recombinant proteins in insect larvae		
JOURNAL	Patent: US 5472858-A 1 05-DEC-1995;		
FEATURES	Location/Qualifiers		
Source	1..1157		
BASE COUNT	212 a 370 c 426 g 149 t		
ORIGIN	/organism="unknown"		
Alignment Scores:			
Pred. No.:	4.2e-82	Length:	1157
Score:	1568.00	Matches:	314
Percent Similarity:	99.05%	Conservative:	0
Best Local Similarity:	99.05%	Mismatches:	3
Query Match:	98.68%	Indels:	0
DB:	6	Gaps:	0
US-09-827-854-16 (1-317) x 115975 (1-1157)			
QY	1	MethylValLeuTrpAlaAlaLeuValThrPheLeuAlaGluCysGlnAlaLysVal	20
Db	62	ATGAAGGTTCTCTGTGGGCTGCGTCTGGTCACTTCTCGGACAGATGCCAGGCCAAGGTG	121
QY	21	GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer	40
Db	122	GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTGGCCAGACAGACAGGAGTGCAGAGC	181
QY	41	GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr	60
Db	182	GGCCGCGCTGGGAACCTGGCACTGGGTGCTTTGGGATTCCTGCGCTGGGTGCAGACA	241
QY	61	LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80
Db	242	CTGTCTGAGCAGGTGCAGAGAGAGCTGCTCAAGTCCCAAGTCAACCAAGAACTGAGGCG	301
QY	81	LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGluGlnLeu	100
Db	302	CTGATGAGCAGAGACCATTAAGAGATTTGAAGGCTTCAAAATCGGAATCGAGGAACAAC	361
QY	101	ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	362	ACCCGCTAGCGGAGAGAACCGGGCACGGCTGTCTCAAGAGAGCTGCACAGCGCGAGGCC	421

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 Db 422 CGCGTGGCGCGGACATGAGGAGCTGTGGCCGCCCTGGTGCATGCTACCGCGGAGGTG 481
 QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 482 CAGGCGATGCTGGCGCAGAGCAGCAGAGAGCTGGCGGTGGCTGGCTCCACTGGCG 541
 QY 161 LysLeuArgGlyArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 Db 542 AAGCGCTTAAGCGGCTCTCCGCGATCCGATCCGAGAACCGCTGGCACTGTAC 601
 QY 181 GlnAlaGlyAlaArgGlyGlnGlyValGlnArgGlyLeuSerAlaIleArgGlyLeuGly 200
 Db 602 CAGCGCGGCGCGCGAGGCGCGCGCGCGCTCAGCGCGCATCCGAGAGCGCTGGGG 661
 QY 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 662 CCCCTGGTGAACAGGCGCGCTGGCGCGCACTGTGGCTCTGGCGCGCGAGCGG 721
 QY 221 LeuGlnGluArgAlaGlnAlaThrGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
 Db 722 CTACAGAGCGGCGCGAGCGCTGGCGCGAGCGCGCGCGATGAGAGATGGGC 781
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
 Db 782 AGTCGAGACCGCGACCGCTGGAGAGGTGAAGAGAGAGGTGGCGAGGTGGCGCCCAAG 841
 QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 Db 842 CTGGAGAGCAGCGCCAGCAGATAGCGCTGCAGGCGCGAGCGCTTCCAGCGCGCGCTCAAG 901
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 Db 902 AGCTGGTTCGAGCCCTGGTGAAGACATGCAGCGCCAGCGCGCGCTGTGTGAGAG 961
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 962 GTGCAGGCTGGCGTGGAGCAGCGCGCGCTGTGGCGCGAGCAATCAC 1012
 RESULT 12
 HUMAPOE
 LOCUS Human apolipoprotein E mRNA, complete cds. linear PRI 08-AUG-1995
 DEFINITION
 ACCESSION M12529
 VERSION M12529.1 GI:178848
 KEYWORDS
 SOURCE Homo sapiens (clone: PHAE112,178,813.) male 57-year old liver cDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS McLean,J.W., Elshourbagy,N.A., Chang,D.J., Mahley,R.W. and Taylor,J.M.
 TITLE Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing of a new variant
 JOURNAL J. Biol. Chem. 259 (10), 6498-6504 (1984)
 MEDLINE 84212473
 PUBMED 6327682
 FEATURES
 source location/Qualifiers
 1..1157
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="19q13.2"
 /clone="PHAE112,178,813."
 /sex="male"
 /tissue_type="liver"
 /dev_stage="57-year old"
 1..1157
 /gene="APOE"
 <1..1157
 mRNA

/gene="APOE"
 /note="APOE mRNA; G00-119-691"
 62..1015
 /gene="APOE"
 /note="precursor"
 /codon_start=1
 /product="apolipoprotein E"
 /protein_id="AA859518.1"
 /db_xref="GI:178849"
 /db_xref="GDB:G00-119-691"
 /translation="MKVLMALLVTFLAGCQAKVEQAVETPEPELRQQTWOSQGRW
 ELALGRFWDLRWVOTLSEVOBELSSOYTOELRALMDETMKLAKVSELEQLRP
 VAETRRARLSKEILOTQARLQADMEYDCGLVYRGVQAMLAGSTPELRVRLASHLR
 KLRRLRDPDDLOKRLAVYQAGAREAGELISIRRELGLPVGQVRAVAVSLAG
 OPIQROANGCERLRARMEKGSRTPRRLDEVRQVAFVRAKLEEQAOQRLQAEARQ
 ARLSKPEPLVEDNQROMAGLIVERVQAAVGTSAAPVPSDNH"
 62..115
 /gene="APOE"
 /note="G00-119-691"
 116..1012
 /gene="APOE"
 /product="apolipoprotein E"
 /note="G00-119-691"
 370
 /gene="APOE"
 /note="a in PHAE112,178]; g in PHAE813; G00-119-691"
 /replace="g"
 410
 /gene="APOE"
 /note="a in PHAE112,178]; g in PHAE813; G00-119-691"
 /replace="g"
 569
 /gene="APOE"
 /note="c in PHAE112,178]; g in PHAE813; G00-119-691"
 /replace="g"
 784
 /gene="APOE"
 /note="t in PHAE112,178]; c in PHAE813; G00-119-691"
 /replace="c"
 859
 /gene="APOE"
 /note="g in PHAE112,178]; a in PHAE813; G00-119-691"
 /replace="a"
 BASE COUNT 212 a 370 c 426 g 149 t
 ORIGIN 30 bp upstream of BstNI site.
 Alignment Scores:
 Pred. No.: 4.2e-82 Length: 1157
 Score: 1568.00 Matches: 314
 Percent Similarity: 99.05% Conservative: 0
 Best local Similarity: 99.05% Mismatches: 3
 Query Match: 98.68% Indels: 0
 DB: 9 Gaps: 0
 US-09-827-854-16 (1-317) x HUMAPOE (1-1157)
 QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 62 ATGAAGGTTCTGTGGGCGCTGTGCTGCTCAATTCTGGCAGGATGCGCAAGGTG 121
 QY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 Db 122 GAGCAGAGCGGTGAGACAGACCGGAGCGCGAGCGCGCAGACCGCAATGGCAGAGC 181
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 Db 182 GCGCAGCGCTGGAACTGGCAGCTGGCTGGCTTTGGGATTCCTGCGCGCTGGGTGAGACA 241
 QY 61 LeuSerGluGlnGlnGlnGluLeuLeuSerSerGlnValTrpGlnGluLeuArgAla 80
 Db 242 CTGCTGAGCAGGTGACAGAGAGCTGCTCAGCTCCCAAGTCACCAAGAACTGAGGCG 301
 QY 81 LeuMetAspGluTrpMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu 100


```

|||||
Db 302 CTGATGACGACGACATGTAAGAGTTGAAGGCTTACAAATCGAATGGAGGACACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 362 ACCCGCGTAGCGAGAGACGCGGGCAGCGCTGTCCAGAGAGTGCAGACGGCGCAGGGC 421
QY 121 ArgLeuGlyAlaAspMetClnuSpValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 422 CGGCTGGCGCGGACATGAGAGACGTGTGCGCGCTGCTGAGTACCGCGCGGAGGTG 481
QY 141 GlnAlaMetLeuGluGlnSerThrGluGluArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCCATCTGCGCCAGAGACCGAGAGCTGGGTGCGCTGCGCTGCCACTGCCCGCC 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 542 AAGCTGCGTAAGCGGCTCTCCGCGCATCCGATGACCTGCAGAGGCGCGTGGCAGTGC 601
QY 181 GlnAlaGlyAlaArgGluGluArgGlyLeuSerAlaAlaArgGluArgLeuGly 200
Db 602 CAGCGCGGGCGCGCGAGCGCGCGCGCTCACGCGCATCCGAGCGCGCTGGGG 661
QY 201 ProLeuValGluGluGlnValArgValaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 662 CCCCTGTGTGAACAGGGCGCGGCGCGCGCACACTGTGGCTCCCTGGCGCGCGCAGCCG 721
QY 221 LeuGlnGluArgAlaGlnAlaThrPglLysGluArgLeuArgAlaArgMetGluGluMetGly 240
Db 722 CTACAGAGAGCGGGCCAGCGCTGGGGCGAGCGGCTGGCGCGCGATGAGAGATGGGG 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValaAlaGluValaArgAlaLys 260
Db 782 ACTCGGACCGCGGACGCGCTGTGACGAGGTGAAGAGCAGTGGCGGAGTGGCGCGCAG 841
QY 261 LeuGluGluGlnAlaGlnGlnThrArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 842 CTGGAGAGACGAGCCAGCAGATACCGCTGCAGGCGCGGCTTCCAGGCGCGCTCAAG 901
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 902 ACTGCTGTGAGACCCCTGTGTGAGACATGACAGCGCATGTGGCGCGCTGTGGAGAG 961
QY 301 ValGlnAlaAlaValaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGCAGGCGCTGTGGCAGCAGCGCGCGCTGTGCGCCAGCAGACATCAGC 1012

RESULT 13
BD004279 1156 bp DNA linear PAT 31-JAN-2002
DEFINITION Apo E humanized mammal.
ACCESSION BD004279
VERSION BD004279.1 GI:18632240
KEYWORDS JP 2001017028-A/3.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1156)
AUTHORS Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.
TITLE Apo E humanized mammal
JOURNAL Patent: JP 2001017028-A 3 23-JAN-2001;
MITSUBISHI CHEMICAL CORP
OS Homo sapiens (human)
PN JP 2001017028-A/3
PD 23-JAN-2001
PF 28-APR-2000 JP 2000128919
PR
PI SHINOBU FUJITA,HIROKI HAMANAKA,YUKO FUKUI,MINESUKE YOKOYAMA PC
A0167/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10,PC
C12N15/09//C07K14/775,
PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,(C12N5/00,C12R1:91) CC

```

```

FEATURES             FH  Key  Location/Qualifiers
source               FT  CDS  Location/Qualifiers
                        1..1156
BASE COUNT          208 a 369 c 432 g 147 t
ORIGIN
Alignment Scores:
Pred. No.:          6,24e-82      Length:      1156
Score:              1565.00      Matches:      315
Percent Similarity: 99.37%      Conservative: 0
Best Local Similarity: 99.37%   Mismatches:  2
Query Match:        98.49%      Indels:      0
DB:                  6           Gaps:        0

US-09-827-854-16 (1-317) x BD004279 (1-1156)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGGTTCTGTGGCTCGTGTGGTGCACATTCCTGGCAGAGATGCCAGGCAAGGTG 120
QY 21 GluGlnAlaValaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAAGCGGTGAAGACAGAGCGGAGCGCGAGCTGCGCACACAGACGAGTGGCAGAGC 180
QY 41 GlyLsnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 181 GGCACAGCGTGGGAACTGGCAGCTGGGCTTTTGGAGTTACTCGCGTGGTGCAGACA 240
QY 61 LeuSerGluGlnValaGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGTCTGACAGAGTGCAGAGAGAGCTGTGCAGCTCCAGGTCCACCAGAACTGAGAGGCG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
Db 301 CTGATGAGACAGACATGAGAGAGTTGAAGGCTTCAAAATCGGAACCTGAGAGAACACTG 360
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 361 ACCCGGTGGCGGAGAGACAGCGGGCAGCGCTGTCCAAAGAGACTCAGCGCGCGCAGGCC 420
QY 121 ArgLeuGlyAlaAspMetClnuSpValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 421 CGGCTGGCGCGGACATGAGAGAGCTGTGGCGCGCGCTGTGTGACATACCGCGCGAGGTG 480
QY 141 GlnAlaMetLeuGluGlnSerThrGluGluArgValaArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTGCGCCAGAGACCGAGAGAGCTGGGTGCGCTGCCCTCCACTGCGCG 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 541 AAGCTGCGTAAGCGGCTCTCCGCGCATGCTGACAGACAGCGCTGCGCACTGATAC 600
QY 181 GlnAlaGlyAlaArgGluGluArgGlyLeuSerAlaAlaArgGluArgLeuGly 200
Db 601 CAGGCGGGGGCGCGGAGAGCGCGCAGCGCGCTGCAGCGCGCATCCGAGAGCGCTGGGG 660
QY 201 ProLeuValGluGlnGluArgValaArgAlaAlaThrValaGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGTGTGAACAGGGCGCGGCTGGCGCGCATCTGTGGCTCCCTGGCGCGCGCAGCCG 720
QY 221 LeuGlnGluArgAlaGlnAlaThrPglLysGluArgLeuArgAlaArgMetGluGluMetGly 240
Db 721 CTACAGAGAGCGGGCCAGCGCTGGGGCGAGCGCGCTGCGCGCGCATGAGAGATGGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValaAlaGluValaArgAlaLys 260
Db 781 AGCCGAGACCGCGGACGCGCTGTGAGAGGTGAAGAGCAGTGGCGGAGTGGCGCGCAG 840
QY 261 LeuGluGluGlnAlaGlnGlnThrArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280

```


Db 841 CTGGAGAGACAGCCAGCATACCGCTGACAGCGCCAGCGCTTCCAGCCCGCCTCAAG 900
Qy 281 SerTTPheGluProLeuValGluAspMetGlnArgGlnTTPAlaGlyLeuValGluIlys 300
|||||
Db 901 AGCTGGTTGAGAGCCCTGGTGGAGAGACATGACAGCGCATATGGCGCGCTGGTGGAGAA 960
Qy 301 ValGlnAlaIaValGlyThrSerAlaIaProValProSerAspAsnHis 317
|||||
Db 961 GTGCAGAGCTGCGGTGGGACACAGCGCGCGCTGTGCCAGGACCAATCAC 1011
RESULT 14
MFAPOE 1178 bp mRNA linear PRI 31-MAR-1995
LOCUS
DEFINITION Monkey mRNA for apolipoprotein E.
ACCESSION X13887
VERSION X13887.1 GI:38054
KEYWORDS apolipoprotein; apolipoprotein E.
SOURCE Macaca fascicularis.
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopitheciinae; Macaca.
REFERENCE 1 (bases 1 to 1178)
AUTHORS Marotti,K.R.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1988) Marotti K.R., The Upjohn Co, 7242 209 713,
301 Henrietta Street, Kalamazoo, MI 49008
2 (bases 1 to 1178)
REFERENCE Marotti,K.R., Whitted,B.E., Castle,C.K., Polites,H.G. and
Melchior,G.W.
TITLE Nucleotide sequence of the cynomolgus monkey apolipoprotein E cDNA
JOURNAL Nucleic Acids Res. 17 (4), 1778 (1989)
MEDLINE 89160349
PUBMED 2922300
COMMENT Data kindly reviewed (20-Mar-1989) by Marotti K.R.
FEATURES
source location/Qualifiers
1..1178
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="1"
/tissue_type="liver"
/clone_lib="lambda gcl1"
83..1036
/codon_start=1
/product="proapolipoprotein E (AA -18 to 299)"
/protein_id="CAA32092.1"
/db_xref="GI:38055"
/db_xref="SWISS-PROT:P10517"
/translation="MKVLMALVTFLAGCOAKYQVPEPEPELRQAEGSGQPM
VLALRFMDYLRWQTLSEVOEELSPVOELTTLMDETMKELKAKSLELEQLSP
VAETFRALSKELQAAQARLGADMEDVSRIVQYSEVOAMLGVEGRVRAFLASLHR
KLRLRLDADDLQRLAVYQAGREGASRLVQSLREGLQPIVEGGRVRAFLASLHR
QPLQRRALGERLQARMEKMSRTRDRLDEYKAEYRAKLEBQAQISLQBARFO
ARKLSWEPELVEDMORWAGLVEKVOAVASTAVPIDNH"
sig_peptide 83..136
mat_peptide 137..1033
misc_feature /product="apolipoprotein E (AA 1 - 299)"
1154..1159
/note="polyA signal"
BASE COUNT 217 a 373 c 426 g 162 t
ORIGIN
Alignment Scores:
Pred. No.: 6.54e-75 Length: 1178
Score: 1443.00 Matches: 295
Percent Similarity: 93.69% Conservative: 2
Best Local Similarity: 93.06% Mismatches: 20
Query Match: 90.81% Indels: 0
DB: 9 Gaps: 0
US-09-827-854-16 (1-317) x MFAPOE (1-1178)
Qy 1 MetLysValLeuTTPAlaIaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20

Db 83 ATGAAGGTTCTGTGGGCTCGTGTGTCACATTCCTGGACAGATGCCAGGCAAGGTG 142
Qy 21 GlnGlnAlaValGlnThrGlnProGlnLeuLeuArgGlnGlnThrGlnTTPGlnSer 40
|||||
Db 143 GACCAACCGGTGGAGCCAGAGACGAGACGACCTTGGCCAGAGAGCTGAGGGCCAGAG 202
Qy 41 GlyGlnArgTTPGlnLeuAlaLeuGlyArgPheTTPAspTyrLeuArgTTPValGlnThr 60
|||||
Db 203 GGCACAGCCTGGGAGCTGGACACTGGGTGCTTTTGGGATTAACCTGGCTGGGGGACAGACA 262
Qy 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
|||||
Db 263 CTCTCTGACAGAGGTGACAGAGAGGAGCTGACAGCCCGCCAGGTTCACCCAGGAACAGAGACG 322
Qy 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerLysLeuGlnGlnGlnLeu 100
|||||
Db 323 CTGATGGACGAGACCATGAAGGATTGAAGGCTTACAAATCGGAACCTGGAGGAACAGCTG 382
Qy 101 ThrProValAlaGlnGlnLeuThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
|||||
Db 383 AGCCCGGTGGCGAGAGAGACGCGGCGACGGCTCTCAAGAGAGCTGCAGCGGCGCAGGCC 442
Qy 121 ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnTTPArgGlyGlnVal 140
|||||
Db 443 CGGCTGGTGGCGACATGAGAGAGACGTGCGACAGCCGCTGGTGCAGTACCGCAGCAGAGT 502
Qy 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 503 CAGGCCATCTGGGCGACAGTACCGAGAGCTGGGGCGGCTCCCTCCACCTGCGC 562
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
|||||
Db 563 AAGCTCGGCAAGGGGCTCTCCCGCATGCTGATGAGACTGCAGAGAGGCGCTGCAGAGTAT 622
Qy 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
|||||
Db 623 CAGGCGGGGCGCGCGAGGGCGCGCGAGCGCGGGGCTCAGCGCCATCCGACGCGCTGGGA 682
Qy 201 ProLeuValGlnGlnGlnLysArgValArgAlaAlaThrValGlnSerLeuAlaGlnPro 220
|||||
Db 683 CCCCTGTGTGAGAGAGGGCGCGGTGGCGGCGCGCTGTGGCTCCCTGGCGACCGCAGCGG 742
Qy 221 LeuGlnGlnArgAlaGlnAlaTTPGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
|||||
Db 743 CTTCAGAGAGCGGGCCAGGCCCTTGGGTGAGCGGCTTCGGCGACGAGTGAAGAGATGGGG 802
Qy 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaIys 260
|||||
Db 803 ACCCGGACCCGCGACCGCTGCGACGAGTGAAGAGACAGGTGGCGGAGGTGCGCGCCAG 862
Qy 261 LeuGlnGlnGlnAlaGlnGlnGlnLeuLeuArgAlaGlnAlaPheGlnAlaArgLeuLys 280
|||||
Db 863 CTGGAGAGACAGCGCCAGAGATTAAGCTTCAGAGCGCGAGCGCTTCCAGGCCCGCCCTCAAG 922
Qy 281 SerTTPheGluProLeuValGluAspMetGlnArgGlnTTPAlaGlyLeuValGluIlys 300
|||||
Db 923 AGCTGGTTGAGAGCCCTGGTGGAGATATGACAGCGCAATGGGCTGGTGGGAGAA 982
Qy 301 ValGlnAlaIaValGlyThrSerAlaIaProValProSerAspAsnHis 317
|||||
Db 983 GTGCAGAGCTGCGGTGGGCGCAGACAGCGCCCTGTGCCATGCACCAATCAC 1033
RESULT 15
AF261279
LOCUS AF261279 5491 bp DNA linear PRI 27-OCT-2000
DEFINITION Homo sapiens apolipoprotein-E gene, complete cds.
ACCESSION AF261279
VERSION AF261279.1 GI:11034800
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 5491)
AUTHORS Nickerson,D.A., Taylor,S.L., Fullerton,S.M., Weiss,K.M.,
Clark,A.G., Stengard,J.H., Salomaa,V., Boerwinkle,E. and Sing,C.F.
TITLE Sequence diversity and large-scale typing of SNPs in the human
JOURNAL Apolipoprotein E gene
MEDLINE Genome Res. 10 (10), 1532-1545 (2000)
PUBMED 20499366
REFERENCE 2 (bases 1 to 5491)
AUTHORS Nickerson,D.A.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2000) Department of Molecular Biotechnology,
University of Washington, Box 357730, Seattle, WA 98195, USA
FEATURES
source 1..5491
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.2"
repeat_region <3..>108
/note="putative"
/rpt_family="MIR"
/rpt_type="dispersed"
variation 73
/frequency="0.01"
/replace="t"
/note="putative"
/rpt_family="MIR"
/rpt_type="dispersed"
variation 308
/frequency="0.01"
/replace="t"
/note="putative"
/rpt_family="MIR"
/rpt_type="dispersed"
satellite <320..>339
/note="putative"
/rpt_type="tandem"
repeat_region <340..>637
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
variation 471
/frequency="0.01"
/replace="g"
variation 545
/frequency="0.01"
/replace="t"
variation 560
/frequency="0.22"
/replace="t"
variation 624
/frequency="0.07"
/replace="c"
satellite <638..>718
/note="putative"
/rpt_type="tandem"
variation 832
/frequency="0.45"
/replace="t"
mRNA join(1060..1094,1855..1920,3013..3205,3786..4645)
/product="Apolipoprotein-E"
variation 1163
/frequency="0.35"
/replace="c"
variation 1522
/frequency="0.01"
/replace="a"
variation 1575
/frequency="0.01"
/replace="t"
CDS join(1878..1920,3013..3205,3786..4503)
/note="APOE"
/codon_start=1

/product="Apolipoprotein-E"
/protein_id="AA627089.1"
/db_xref="GI:11034801"
/translation="MKVLAALLVPLAGCAKVEQAVETEPPELRROOTEMOSGORM
ELAGFPMYLRWQTLSPNOEELISSQVQELRALMETKEKAYSELEDTLP
VAEETRLRSKELOAARLADMEVRCRLVQYREVOAMLGOSTEELRVLASLR
KIKRRLRLDADIDQRLAVYQAGARGARGLSALREKGPLVEQGRVAATVGSILAG
QPIQERAAQMGRLRAVMEGSRTRDRIDEVKEDVAEVRARLEQDAQDIRLQAEAFQ
ARLKSFEPLVEDMQRMAGLVEKVOAAVGTSAAPVPSDNH"
variation 1998
/frequency="0.10"
/replace="a"
repeat_region <2124..>2435
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
variation 2440
/frequency="0.21"
/replace="a"
/replace="a"
repeat_region <2569..>2848
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
variation 2907
/frequency="0.01"
/replace="g"
variation 3106
/frequency="0.01"
/replace="c"
repeat_region <3472..>3588
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
variation 3673
/frequency="0.01"
/replace="g"
variation 3701
/frequency="0.01"
/replace="ct"
variation 3937
/frequency="0.14"
/replace="c"
variation 4036
/frequency="0.01"
/replace="t"
variation 4075
/frequency="0.07"
/replace="t"
repeat_region <4735..>5056
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
variation <5065..>5476
/note="putative"
/rpt_type="dispersed"
variation 5229
/frequency="0.03"
/replace="gg"
variation 5229
/frequency="0.07"
/replace="t"
variation 5229
/frequency="0.40"
/replace="g"
variation 5230
/frequency="0.13"
/replace="a"
variation 5361
/frequency="0.06"
/replace="c"
BASE COUNT 1029 a 1663 c 1603 g 1196 t
ORIGIN

Alignment Scores:

Pred. No.:	4,32e-71	Length:	5491
Score:	1388.50	Matches:	303
Percent Similarity:	60.608	Conservative:	0
Best Local Similarity:	60.608	Mismatches:	3
Query Match:	87.38%	Indels:	194
DB:	9	Gaps:	1

US-09-827-854-16 (1-317) x AF261279 (1-5491)

```
QY 12 PheLeuAlaGlyCysGlnAlaIysValGluGlnAlaValGluThrGluProGluProGlu 31
    |||
DB 3003 TTCACACAGAGATGACGAGCCCAAGGTGGAGAGCAAGCGGTGGAGACAGAGCCGAGCCGAG 3062
QY 32 LeuArgGluGlnThrGluThrGluThrGluThrGluThrGluThrGluThrGluThrGluThr 51
    |||
DB 3063 CTGCGCCAGACAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3122
QY 52 TrpAspTyrLeuArgTyrPylGlnThrLeuSerGluGlnValGlnGlnGlnLeuLeuSer 71
    |||
DB 3123 TGGGATTACCTGGCTGGGTGGAGACACGTCTGAGCAGAGTGGAGAGGAGGAGGAGGAGGAGGAG 3182
QY 72 SerGlnValThrGlnGlnLeu----- 78
    |||
DB 3183 TCCACAGTACCCAGAGACTGAGGTGAGTGCCTCCATCCCTGGCCCTTGACCCCTCGTG 3242
QY 78 ----- 78
DB 3243 GCGGCTATACCTCCCAAGGTCCAGGTTTTCATTCGCCCCCTGCTGCTTGGGGG 3302
QY 78 ----- 78
DB 3303 CCTGGSTCTCTGCTGCTTCTAGCTTCCCTTCCCATTTCTGACTCCTGCTTACTCTTC 3362
QY 78 ----- 78
DB 3363 TGGAAATTCCTCTCTCAGCTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3422
QY 78 ----- 78
DB 3423 CGTCTGAGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3482
QY 78 ----- 78
DB 3483 CTCACGTGTGTGCGCCAGGCTGTGTGAACCTCTGGGCTCAAGCGATCCTCCCGCTCGG 3542
QY 78 ----- 78
DB 3543 CTTCCCAAGTGTGGGATTAGAGGATGAGCCACCTTGCCGAGCTCTAGCTCTCTCT 3602
QY 78 ----- 78
DB 3603 TGTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 3662
QY 78 ----- 78
DB 3663 GCGTGTGCGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3722
QY 78 ----- 78
DB 3723 CCCCATCCAGCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3782
QY 79 -ArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGln 98
    |||
DB 3783 CAGGGGCTGATGAGAGACCATGAAGAGTTGAAGGCTTACAAATCGAACTGGAGGA 3842
QY 98 uGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAl 118
    |||
DB 3843 ACAACTGACCCCGGTGGCGGAGAGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3902
QY 118 aGlnAlaArgLeuGlnAlaAspMetGluAspValCysGlnArgLeuValGlnTyrArgGln 138
    |||
```

```
DB 3903 GCAGGCCCGGCTGGGCGCGAGACATGGAGACGTGTGCGGCCGCTGTGTCAGTACCGCGG 3962
QY 138 yGluValGlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerH 158
    |||
DB 3963 CGAGGTGACAGGCCATGCTGTGGCCAGAGACCCGAGAGGCTGGCGGTGCTGCTGCCA 4022
QY 158 sLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAl 178
    |||
DB 4023 CCTGGCCAGAGCTGCGTAAAGGCGCTCTCCGCGATGCGCATGACGTGACAAAGCGCTGGC 4082
QY 178 aValTyrGlnAlaGlyAlaArgGluGlnValAlaGluArgGlyLeuSerAlaIleArgGluAr 198
    |||
DB 4083 AGTGTACCAAGCGCGGCGCGGAGGCGCCGAGCGCGGCTTCAGCGCCATCCGCGAGCG 4142
QY 198 gLeuGlyProLeuValGluGlnGlnValArgValArgAlaIleThrValGlySerLeuAlaG 218
    |||
DB 4143 CTTGGGGCCCTGTGTGGAACAGGCGCGTGTGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTG 4202
QY 218 yGlnProLeuGlnGluArgAlaGlnAlaIleThrPylGluArgLeuArgAlaArgMetGluG 238
    |||
DB 4203 CCAGCGGCTACAGAGAGCGGCGCCAGGCTGGGGCGAGCGGCTGGCGCGGATGGAGGA 4262
QY 238 uMetGlySerArgTyrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAr 258
    |||
DB 4263 GATGGGACCGCGACCGCGGACCGGCTGTGAGCGAGGTGAAGAGAGAGGTGGAGGTGGC 4322
QY 258 gAlaLysLeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaAr 278
    |||
DB 4323 GCGCAAGCTGGAGAGAGAGCCAGACAGATACGCTGCGAGCGCGAGCCCTTCCAGGCGCG 4382
QY 278 gLeuLysSerTyrPheGluProLeuValGluAspMetGlnArgGlnTyrPalagLysLeuVa 298
    |||
DB 4383 CCTCAAGAGCTGTGTGAGCCCTGTGTGAAGACATGACAGCGCCAGTGGCGCGGCTGTGT 4442
QY 298 lGluLysValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
    |||
DB 4443 GGAGAGGTGACAGGCTGCGTGGGACACAGCCGCGGCTGTGTGCCAGGACATATCAC 4500
```

Search completed: March 14, 2003, 17:32:29
Job time : 2314.23 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:08:17 ; Search time 178.728 Seconds
(without alignments)
3994.237 Million cell updates/sec

Title: US-09-827-854-16
Perfect score: 1589
Sequence: 1 MKVLMALLVFLAGCOAKV.....VEKVOAAGTSAPVPSDNH 317

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlp
-O=/cgn2.1/USPRO.spool/us09827854/runat_11032003.101609_27476/app_query.fasta.1.3576
-DB=Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCCALIGN=200 -THR_SCORE=0 -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MOD=LOCAL -OUTFMT=pro -NOR=ext -HEAPSIZE=500 -MILEN=0 -MAXLEN=2000000000
-USER=us09827854.cgn.1.1.1201 @runat_11032003.101609_27476 -NCPU=6 -ICPU=3
-NO_XLPYX -NO_MMAB -LARGQUERY -NEC_SCORES=0 -WAIT -LONGLOG -DEV=TIMEOUT=120
-WALLTIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
N.Geneseq_101002: *
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1589	100.0	1156	22	AAF84314	Human ApoE2 coding
2	1589	100.0	1156	24	AAD22049	Human apolipoprotein
3	1582	99.6	1156	24	AAD22050	Human apolipoprotein
4	1577	99.2	954	24	AAD26035	Human apolipoprotein
5	1577	99.2	1110	7	AAH60409	Human apolipoprotein
6	1577	99.2	1147	22	ABA83113	Apolipoprotein E o
7	1577	99.2	1156	22	AAF84315	Human apolipoprotein
8	1577	99.0	1156	24	AAD22048	Human apolipoprotein
9	1573	99.0	1156	24	AAD22052	Human apolipoprotein
10	1569	98.7	1156	24	AAD22051	Human apolipoprotein
11	1568	98.7	1157	17	AAT06957	Human apolipoprotein
12	1568	98.7	1157	24	ABN95746	Human apolipoprotein
13	1568	98.7	1157	24	ABK64514	Gene #2244 used to
14	1568	98.7	1157	24	ABL65450	Human benign prost
15	1567	98.6	1110	6	AAH50450	Lung cancer relate
16	1565	98.5	1156	22	AAF84316	Sequence encoding
17	1565	98.5	1156	24	AAD22047	Human ApoE4 coding
18	1473.5	92.7	1279	22	AA522437	Human apolipoprotein
19	1454.5	91.5	1107	19	AAK75756	Human apolipoprotein
20	1388.5	87.4	9360	24	ABL31915	Human apolipoprotein
21	1388.5	87.4	10716	24	AAD26034	Human apolipoprotein
22	1383	87.0	3805	20	AA209524	Human Apo E genomic
23	1383	87.0	3805	20	AA209526	Human Apo E genomic
24	1368.5	86.1	10716	24	AAD26108	Human apolipoprotein
25	1155.5	72.7	965	24	AAD32081	Bovine ApoE gene.
26	1155.5	72.7	5617	24	AAD32077	Human alpha-1-anti
27	1155.5	72.7	6026	24	AAD32075	Human alpha-1-anti
28	1130	71.1	1126	19	AAV29159	Human albumin prom
29	977	61.5	936	15	AAQ69101	Nucleotide sequenc
30	977	61.5	936	17	AAT18070	ApoE4Lx2 protease
31	976	61.4	660	18	AAT69792	Human ApoE4Lx2 cdn
32	939	59.1	597	17	AAT18068	Partial human apo
33	930	58.5	597	15	AAQ69096	Human ApoE4 cdna.
34	876	55.1	1381	22	AA522673	ApoE4L protease cd
35	859.5	54.1	786	21	AAF18114	Human cdna encodin
36	670	42.2	600	20	AAH98595	Lung cancer associ
37	651	41.0	407	24	ABK34238	EST clone CP147.
38	625	39.3	478	24	ABN95994	Human cdna for nov
39	625	39.3	478	24	ABL62679	Gene #2492 used to
40	625	39.3	478	24	ABL67340	Colon adenocarcino
41	594	37.4	499	22	AA503049	Thyroid cancer rel
42	506	31.8	330	12	AAQ11980	Human diagnostic a
43	504	31.7	405	21	AAQ02139	Human apolipoprote
44	485	30.5	345	22	AAH98479	Human secreted pro
45	413	26.0	253	19	AAV39612	Human EST-derived
						E2 allele DNA enco

ALIGNMENTS

RESULT 1
AAF84314
AAF84314 standard; CDNA: 1156 BP.
AAAF84314;
21-JUN-2001 (first entry)
Human ApoE2 coding sequence.
Human; ApoE2; Alzheimer's disease; arteriosclerosis; ss.
Homo sapiens.
Key CDS
Location/Qualifiers
61..1014
/*tag= a
/product= "Human ApoE2"

XX JP2001017028-A.
 PN
 XX 23-JAN-2001.
 PD
 XX
 PF 28-APR-2000; 2000JP-0128919.
 PR 06-MAY-1999; 99JP-0125647.
 PA
 XX (MITU) MITSUBISHI CHEM CORP.
 DR WPI: 2001-285406/30.
 DR P-PSDB: AAB80996.
 XX
 XX New apoe humanized mammalian cell useful for screening for agents
 PT useful for treating or preventing Alzheimer's disease and
 PR arteriosclerosis -
 PS Disclosure; Page 11-12; 22pp; Japanese.
 CC The present invention relates to an APOE humanised mammalian cell. The
 CC present sequence is the coding sequence for human APOE2, which was used
 CC in the method of the present invention. The APOE humanised mammalian cell
 CC can be used for screening for agents useful for treating or preventing
 CC Alzheimer's disease and arteriosclerosis.
 CC
 XX Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;

Alignment Scores:
 Pred. No.: 5 29e-116 Length: 1156
 Score: 1589.00 Matches: 317
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-827-854-16 (1-317) x AAF84314 (1-1156)

QY 1 MetLysValLeuTrpAlaIleuLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
 DB 61 ATGAAGTTCTGTGGGCTGTGGTGTGTCACATTCCTGGCAGAGATGCCAGGCAAGG 120
 QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 121 GAGCAAGCGGTGAGAGACAGAGCGGAGCCGAGCTCGCAGCAGACCGAGTGCGAGAGC 180
 QY 41 GlyGlnArgTrpGluLeuAlaGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 60
 DB 181 GGGCAGCGGTGGAGTGGACTGGGTGCTTTGGGATTACCTGGCGTGCGAGACA 240
 QY 61 LeuSerGluGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 DB 241 CTGTCTGACAGAGTGCAGGAGGAGCTGTCTGAGTCCAGAGTCCAGAGAACAGAGGCG 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
 DB 301 CTGATGACGAGACCAATGAAGGAGTTGAAGGCTTACCAATCGGAACAGAGCAACTG 360
 QY 101 ThrProValAlaGlnGluThrArgAlaArgLeuSerGlnLeuGlnAlaGlnAla 120
 DB 361 ACCCGCGGTGGAGAGAGCGCGGACGCTGTCCAGAGAGTGCAGCGCGGCGAGGCC 420
 QY 121 ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 DB 421 CGGCTGGCGCGGACACTGAGAGACGTGTGCGGCGCTGTGTGAGTACCGCGCGAGGGG 480
 QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 481 CAGGCCATGCTGGCCAGAGACACGAGAGAGTGGCGGCTCGCTCCCTCCACCTGCGCC 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 DB 541 AAGCTGCGTAAGCGGCTCTCCGCGATGCGGATGACCTTCAGAAAGTGCTGGAGAGTAC 600

QY 181 GlnAlaGlyAlaArgGluGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 601 CAGCGCGGGGCCCGGAGAGGCGCGGAGCGGCTCTGAGCCATCCGCGAGCGCTGGGG 660
 QY 201 ProLeuValGlnGlnGlnLysArgValArgAlaIleThrValGlySerLeuAlaGlyGlnPro 220
 DB 661 CCCCTGTGGAAACAGGGCGCGGTGGGCGGCTGTGGGCTCTGTGGCGGCGAGCCG 720
 QY 221 LeuGlnGluArgAlaGlnAlaIleArgGlyGlyArgLeuArgAlaArgMetGluGlnMetGly 240
 DB 721 CTACAGAGAGCGGGCCAGGCTGTGGGCGAGCGGCTGTGGCGGATGAGAGATGGGG 780
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
 DB 781 AGCCGAGACCCGCGACCGCTGTGACGAGGTGAAGAGACAGTGGCGGAGGTGGCGCCCAAG 840
 QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnIleArgLeuLys 280
 DB 841 CTGGAGAGACAGGCCAGCAGATACGCTGACAGGCCGCTTCCAGGCCGCGCTCAG 900
 QY 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 DB 901 AGCTGTGTGAGAGCCCTGTGTGGAAGACATGACAGCGCAGTGGGCGGCTGTGTGAGAAAG 960
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 961 GTGCAGGCTGCGGTGGGACACAGCGCGCCCTGTGCTCCAGCAGCAATCAGC 1011

RESULT 2
 AAD22049
 ID AAD22049 standard; DNA; 1156 BP.

AC AAD22049;

DT 12-FEB-2002 (first entry)

XX Human apolipoprotein E (apoE) isoprotein, apoE2 DNA.

KW Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;

OS hypertriglyceridaemia; low density lipoprotein; LDL; ds.

XX Homo sapiens.

FT sig_peptide Location/Qualifiers

FT CDS 61..1014

FT /tag= a /product= "Human apoE isoprotein, apoE2"

FT mat_peptide 61..114

FT /tag= b

FT /tag= c

FT /product= "Mature human apoE isoprotein, apoE2"

XX WO200177136-A1.

PD 18-OCT-2001.

PF 06-APR-2001; 2001WO-US11358.

XX 06-APR-2000; 2000US-0544386.

PR 04-OCT-2000; 2000US-0679088.

PR 05-APR-2001; 2001US-0827854.

XX (KOSP-) KOS PHARM INC.

PA (UYBO-) UNIV BOSTON.

XX Zannis VI, Kypros KE;

XX WPI: 2002-010885/01.

DR P-PSDB: AAE13295.

XX New apolipoprotein E polypeptide and nucleic acid, useful for lowering

221 LeuGInGuaRgaLaGInaATIpGlyGuaRgLeuARgaLaARgmetGluGluGluGly 240

CC normally functioning apoe gene or low density lipoprotein (LDL)

CC for haplotyping and/or genotyping the APOE gene. The haplotyping
CC methods of the invention are useful for improving the efficacy and
CC reliability of several steps in the discovery and development of
CC drugs for treating diseases associated with APOE activity, e.g.
CC familial dysbetalipoproteinemia, type III hyperlipoproteinemia,
CC atherosclerosis, and Alzheimer's disease. They are useful to validate
CC APOE as a candidate agent for treating a specific condition or disease
CC predicted to be associated with APOE activity and in the design of
CC clinical trials of candidate drugs for treating a specific condition
CC or disease predicted to be associated with APOE activity. Genotyping
CC or haplotyping methods are useful to screen for compounds targeting
CC APOE to treat a specific condition or disease associated with APOE
CC activity. The present sequence is a cDNA encoding human APOE protein.
CC APOE gene is located on chromosome 19q13.2.

Sequence 954 BP; 168 A; 290 C; 377 G; 119 T; 0 other;

Alignment Scores:

Pred. No.:	3,73e-115	Length:	954
Score:	1577.00	Matches:	316
Percent Similarity:	99.68%	Conservative:	0
Best local Similarity:	99.68%	Mismatches:	1
Query Match:	99.24%	Indels:	0
DB:	24	Gaps:	0

US-09-827-854-16 (1-317) x AAD26035 (1-954)

QY	1	MeuysValLeu	rrpLaalaleu	LeuValThr	Phe	Leu	laGlyCys	Gln	laIaIys	Val	20											
Db	1	ATGAAGTCTCTG	TGGCTGGCTTCTG	TGTCACATCTCTG	CAAGATCC	AGGCCA	AGGTG				60											
QY	21	Glu	Gln	laIaVal	Glu	rrpGluPro	Glu	rrpGlu	Leu	rrgGln	Thr	40										
Db	61	GAGCAACGGGT	GGAGACAGAGCCG	AGCCGAGCTGG	CCACAGACAG	CACTGG	CAGACC				120											
QY	41	Gly	Gln	Ar	grrpGlu	Leu	laIaLeu	Gly	Ar	Phe	rrpAsp	rrpLeu	rrgrrpVal	Gln	Thr	60						
Db	121	GGCCAGCGCTGG	GAACAGGCACTGG	GTCTCGCTTTGG	ATTAACCGCTGG	CTGGTGC	ACAACA									180						
QY	61	Leu	Ser	Glu	Gln	rrpValGln	Glu	rrpLeu	Ser	Gln	rrpValThr	Gln	rrpGlu	Leu	rrpGlu	rrpVal	Ala	80				
Db	181	CTGTCTACGAG	GTGCAGAGAGAGCTCT	CAAGCTCC	CAAGGTAC	CCAGGAAC	CTAGAGGCG											240				
QY	81	Leu	Met	Asp	Glu	rrpThr	Met	Gly	Glu	rrpLeu	Gly	Ala	Ar	rrpCys	Ser	Glu	rrpLeu	Gln	Glu	rrpGln	Leu	100
Db	241	CTGATGACG	ACGACCATGAAGAG	ATTGAAGGCT	ACAAATCGGAAC	CTGGAGGA	CAAACTG															300
QY	101	Thr	Pro	Val	laGlu	Glu	rrpThr	rrpVal	Ar	Glu	rrpLeu	Ser	Gly	Ser	Glu	rrpLeu	Gln	laIa	laGln	laIa	120	
Db	301	ACCCGGGTGG	CGGAGGAGCGGGCG	CTGTCCAAGAG	GTGCGAGCGCG	CAAGCC																360
QY	121	Arg	rrpLeu	Gly	laAsp	rrpMet	Gly	rrpAsp	Val	Cys	Gly	Ar	Glu	rrpVal	Gln	rrprrpArg	Gly	rrpVal	140			
Db	361	CGGCTGGG	CGCGGACATGAGAG	AGTGTGGCGCGCTGT	GTGCATGCCG	CGGAGGTG																420
QY	141	Gln	la	Met	Leu	Gln	Gly	Gln	Ser	Thr	Gly	Gln	Leu	rrpVal	Ar	Glu	rrpLeu	laSer	His	Leu	rrpArg	160
Db	421	CAGGCCAT	GTCTGGGCGAGAC	ACCGAGAGACTGG	GGTGGCTCGCTCC	CACTGCGC																480
QY	161	Lys	Leu	rrpArg	Lys	rrpLeu	rrpLeu	rrpAsp	laAsp	rrpAsp	Leu	Gln	rrpLys	Cys	Leu	laIa	Val	Tyr	180			
Db	481	AAGTGGT	GAAGCGGCTCTCCG	GTAGCCGATGACTCGA	GAAGCGCTCG	CAAGTAC																540
QY	181	Gln	la	laIa	Gly	laAr	rrpGlu	Gly	rrpGlu	rrpVal	Leu	Ser	rrpAla	Leu	rrpArg	rrpVal	Glu	Gly	200			
Db	541	CAGCGCGGG	CGCGAGGGCGCGAG	CGCGGCTCA	CGCCCAATCCG	CGAGCGCTCGGG																600
QY	201	Pro	Leu	rrpVal	Gln	Gln	rrpArg	rrpVal	Ar	Glu	rrpAla	laThr	rrpVal	Lys	Ser	rrpLeu	laGly	Gln	rrpPro	220		
Db	601	CCCTGGT	GAACAGGGCGCGT	GGGGCGGCAC	TTGGGCTCTCGG	CGGGCAAGCG																660
QY	221	Leu	Gln	Glu	rrpVal	Gln	laIa	rrprrpGly	Glu	rrpVal	rrpLeu	rrpVal	laAr	rrpMet	Gln	rrpGlu	rrpMet	Gly	240			

Db	661	CTACAGAGAGCGGGCCAGAGCCTGGGGGGAGAGCGCGTGGCGCGCGAGATGAGAGAGATGGC	720
Qy	241	SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys	260
Db	721	AGCCGGACCCCGACCGCTGTGAGAGAGTGAAGAGCAGGTGGCGGAGTGGCGCCACG	780
Qy	261	LeuGluGluGlnIleArgLeuGlnIleArgLeuGlnIleArgLeuGlnIleArgLeuLys	280
Db	781	CTGGAGGAGCGAGGCCACGACATACGCTTGCAGAGCCGACAGGCGCTTCCAGAGCCCGCTTAAG	840
Qy	281	SerTrpPheGluProLeuValGluAspMetGlnArgGlnIleArgLeuValGluLys	300
Db	841	AGCTGGTTGAGCCCTGTGTGGAAGACATGACGCCACAGTGGGCGGGGTGTGTGAGAAC	900
Qy	301	ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis	317
Db	901	GTGGAGGCTGCCGTGGGACACAGGCGCCCTCTGTGCCACGCAATTCAC	951

RESULT 5
AN60409

ID	standard; DNA; 1110 BP
AAN60409	

AC AAN60409;

DT 01-JAN-1980 (first entry)

Human apolipoprotein-E.

KW Apolipoprotein-E; hyperlipidemia; arteriosclerosis; ss

OS Homo sapiens.

aa	Key	Location/qualifiers
FH		

	LD...J06
CDS	/ *tag= a
EI	
ET	

/*tag= a

PN AU8547513-A.

PD 24-APR-1986.

PF 17-SEP-1985; 85AU-0047513.

11-JUN-1985; 85JP-0126989.

PA (MITU) MITSUBISHI CHEM IND KK.

DR WPI; 1986-150217/24.

DR P-PSDB; AAP60507.

PT New DNA sequence coding for human apolipoprotein-E - and

XX

XX
XX

CC apolipoprotein-E (or who produce abnormal forms of this molecule)
CC
CC the enzyme protein is used to treat subjects who are deficient in
CC

888

XX 3

Sequence 1110 BP; 198 A; 354 C; 415 G; 143 T; 0 other;

Alignment Scores:

Pred. No.:	4,43e-115	Length:	1110
Score:	1577.00	Matches:	316
Percent Similarity:	99.68%	Conservative:	0
Best Local Similarity:	99.68%	Mismatches:	168
Query Match:	99.24%	Indels:	0
DB:	7	Gaps:	0

US-09-827-854-16 (1-317) x AAN60409 (1-1110)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20

DB 15 ATGAAAGTTCTGTGGCTGTGGCTGTGCACATTCCTGGCAGAGATCCAGGCGCAAGGCTG 74
 QY 21 GUGAAlaValaGluThrGluUpProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 75 GAGCAAGCGGTGGAGACAGACGCCGAGCCGACCTCGCCAGCAGACCGATGGCGAGAGC 134
 QY 41 GUGlInaGTrpPoluLeuAlaLeuGluGlyArgPheTrpAspTyrLeuArgTrpValaGlnThr 60
 DB 135 GGGCAGCGGTGGGAACTGGGACTGGGTGGCTTTGGGATTACTGGCGGTGGTGCAGACA 194
 QY 61 LeuSerGluGlnValaGlnGluGluLeuSerSerGlnValaThrGlnGluLeuAlaGla 80
 DB 195 CTGTCTGACGACGAGTGCAGAGAGCTGCTCAGCTCCAGGTCCACGAGAACTGAGAGCGC 254
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
 DB 255 CTGATGAGACAGACCATGAAGAGAGTGAAGCCTACAAATCGGAACCTGAGAGAACAACTG 314
 QY 101 ThrProValaAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 315 ACCCCGCTGGCGGAGAGACGCGGCGACGCTGTCCAAGAGCTGCAGCGCGCGCAGGCC 374
 QY 121 ArgLeuGluValaAspMetGluAspValaLysGlyArgLeuValaGlnThrArgGluVala 140
 DB 375 CGGCTGGGCGCGACATGGAGAGAGTGTGCGCGCTGTGTGATGATACCGCGCGAGAGTGG 434
 QY 141 GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValaArgLeuAlaSerHisLeuArg 160
 DB 435 CAGGCGATGCTGGCGCCAGAGACCGAGAGAGCTCGGGTGGCTCCCTCCACCTCGCGC 494
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValaTyr 180
 DB 495 AAGCTCGTAAAGCGCTCTCCCGCATGCGCATGACCTCGAGAAGCGCTGCAGAGTAC 554
 QY 181 GlnAlaGluValaArgGluGluGluValaArgGlyLeuSerAlaHisArgGluValaGluGly 200
 DB 555 CAGGCGGGGCGCGCAGAGGCGCGAGCGCGCTCAGCGCCATCCGCGAGCGCTGGGG 614
 QY 201 ProLeuValaGluGlnGluArgValaArgAlaAlaThrValaGlySerLeuAlaGlyGlnPro 220
 DB 615 CCCTGGTGGAAACAGGCGCGCTGCGGCGCGCAGCTGTGGCTCCCTGGCGCGCGCAGCGC 674
 QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
 DB 675 CTACAGAGAGCGGCGCCAGGCGCTGGGCGAGCGCGCTGCGCGCGGATGGAGATGGGC 734
 QY 241 SerArgThrArgAspArgLeuAspGluValaLysGluGlnValaAlaGluValaArgAlaLys 260
 DB 735 AGCGGAGACCGCGACCGCTGGCGAGGAGTGAAGAGCAGGTGGCGAGGTGGCGCGCAG 794
 QY 261 LeuGluGluGlnAlaGlnGlnLeuArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 DB 795 CTGAGAGAGAGCGCCAGCAGATACCTGACGCGCAGCGCGCTTCACAGCGCGCCCTCAAG 854
 QY 281 SerTrpPheGluProLeuValaGluAspMetGlnArgGlnTrpAlaGlyLeuValaGluLys 300
 DB 855 AGCTGTGTGAGAGCCCTGTGGGAAAGACATGCACCGCAGTGGGCGCGGTGTGGAGAG 914
 QY 301 ValGlnAlaAlaValaGlyThrSerAlaAlaProValaProSerAspAsnHis 317
 DB 915 GTGCAAGGCTGCGCTGGGCGACAGCGCGCCCTGTGCTGCCACGACATATCAC 965

RESULT 6
 ABA83113
 ID ABA83113 standard: DNA; 1147 BP.

AC ABA83113;
 DT 08-FEB-2002 (first entry)
 DE Apolipoprotein E ovarian tumour marker gene, SEQ ID NO:63.

KW Ovarian tumour marker gene; human; overexpression; upregulation;
 KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;
 KW identification; serous cystadenoma; borderline serous tumour;
 KW serous cystadenocarcinoma; mucinous cystadenocarcinoma;
 KW mucinous cystadenoma; borderline mucinous tumour; endometrioid carcinoma;
 KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;
 KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;
 KW immune response pathway; cell proliferation regulation; protein folding;
 KW membrane localised; secreted; therapeutic target; cytostatic;
 KW gene therapy; vaccine; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200175177-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 03-APR-2001: 2001WO-US10947.
 XX
 PR 03-APR-2000; 2000US-194336P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Morin RJ, Sherman-Baust CA, Pizer ES, Hough CD;
 XX WPI: 2001-626450/72.
 DR P-PSDB: ABB50287.
 XX
 PT Detecting and identifying ovarian tumor, identifying increased risk for
 PT developing ovarian cancer, and determining effectiveness of ovarian
 PT cancer treatment, by measuring expression level of ovarian tumor marker
 PT gene -
 XX
 PS Claim 23; Page 105-106; 140pp; English.
 XX
 CC The invention relates to methods for diagnosing and prognosing ovarian
 CC tumors in an individual via the detection and measurement of the
 CC expression of ovarian tumour marker genes (ABA83081-ABA83122, ABA83180,
 CC ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179,
 CC ABA83181 and ABA83183). The methods of the invention are useful for
 CC detecting an ovarian tumour in a patient, for identifying an individual
 CC at increased risk for developing ovarian cancer. In prognostic tests for
 CC assessing the relative severity of ovarian cancer, in tests for
 CC monitoring a patient in remission from ovarian cancer and in tests for
 CC monitoring disease status in a patient being treated for ovarian cancer.
 CC The methods can additionally be used to identify a particular tumour as
 CC being an ovarian tumour (i.e., an epithelial ovarian tumour selected from
 CC serous cystadenoma, borderline serous tumour, serous cystadenocarcinoma,
 CC mucinous cystadenoma, borderline mucinous tumour, mucinous
 CC cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma,
 CC clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner
 CC tumour. The ovarian tumour marker genes of the invention were identified
 CC using SAGE (serial analysis of gene expression) and were found to be
 CC overexpressed in a broad variety of ovarian epithelial tumour cells
 CC relative to normal ovarian epithelial cells. The marker genes are
 CC implicated in immune response pathways, in the regulation of cell
 CC proliferation and in protein folding, and many of these are membrane-
 CC localised or secreted. In addition to their use as diagnostic and
 CC prognostic markers, the ovarian tumour marker genes or their encoded
 CC proteins may be used as therapeutic targets for the treatment and
 CC prevention of ovarian cancer. Sequences ABA83081-ABA83122, ABA83180,
 CC ABA83182 and ABA83184 represent the ovarian tumour marker genes of
 CC the invention.
 CC
 SQ Sequence 1147 BP; 210 A; 365 C; 425 G; 147 T; 0 other;
 XX
 XX
 Alignment Scores:
 Pred. No.: 4.6e-115 Length: 1147
 Score: 1577.00 Matches: 316
 Percent Similarity: 99.68% Conservative: 0
 Best Local Similarity: 99.68% Mismatches: 1
 Query Match: 99.24% Indels: 0
 DB: 22 Gaps: 0


```

US-09-827-854-16 (1-317) x AAF83113 (1-1147)
OY      1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
        |||
        46 ATGAAGGTTCTGTGGCGCTGGCTGGTGCACATTCTCTGGCAGAGATGCCAGGCCAAGGTG 105
OY      21 GluGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
        |||
        106 GAGCAAGCGGTGAGACAGAGCCGAGCTGCCGACAGACAGACGAGTGGCAGAGC 165
OY      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
        |||
        166 GGCACGCGCTGGGAACCTGGCACTGGTGGCTTTGGATTACTCTGCTGGGTGGCAGACA 225
OY      61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
        |||
        226 CTGTGTGACAGAGTGCAGAGAGAGCTGCTCACTCCAGATCACCAGAACTGAGGGCG 285
OY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
        |||
        286 CTGATGACAGACCATGACATGAGAGTGAAGGCCCTCAATCGGAAGCTGGAGACACACTG 345
OY      101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
        |||
        346 ACCCGGTGGCGAGAGACAGCGGCGACGCTGCCAAGAGACTCAGAGCGCGCCAGAGCC 405
OY      121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
        |||
        406 CGGCTGGGCGCGACATGAGAGACTGTGGCGCCGCTGTGTCACTACCGCGCCAGAGTG 465
OY      141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
        |||
        466 CAGGCGATGCTGGCGCAGAGACCGAGAGAGCTGGCGGTGGCGCTCCACCTGGGCG 525
OY      161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
        |||
        526 AAGCTGCGTAAAGCGGCTCTCCGCGATGCCGATGACGCAAGAACGCGTGGCACTGTAC 585
OY      181 GlnAlaGlyAlaArgGlnGluGluGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
        |||
        586 CAGCGCGGGGCGCGCGAGGGCGCGCGCTGCACGCGCCATCCGCGCGCGCTGGGGG 645
OY      201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
        |||
        646 CCCCTGGTGGAAACAGAGCGCGGTGGCGCGCCACTGTGGGTCTGGCGCGCCAGCGG 705
OY      221 LeuGlnGluArgAlaGlnAlaThrGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
        |||
        706 CTACAGGAGCGGGCCAGCGCTGGGGCGAGCGGCTGGCGCGCGGATGAGAGATGGGCG 765
OY      241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGlnValArgAlaLys 260
        |||
        766 AGCCGAGACCGCGACCGCTGGACAGGTGAGAGAGAGGTGGCGAGGTGGCGCCCAAG 825
OY      261 LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
        |||
        826 CTGGAGAGAGAGCGCCACAGATAGCTCGAGCGCGCGGCTTCAGAGCCCGCTCAAG 885
OY      281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
        |||
        886 AGCTGGTTCGAGCCCTGCTGGTGAACATGACAGCCGAGTGGCGCGCTGTGTGAGAGAG 945
OY      301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
        |||
        946 GTGCAGGCTCGCGTGGGACCAACGCGCCCTGTGTGCCAGGACAAATCAC 996

```

```

RESULT 7
AAF84315
ID AAF84315 standard; cDNA; 1156 BP.
XX
AC AAF84315;
XX

```

```

DT      21-JUN-2001 (first entry)
XX
DE      Human ApoE3 coding sequence.
XX
KW      Human; ApoE3; Alzheimer's disease; arteriosclerosis; ss.
XX
OS      Homo sapiens.
XX
FH      Key
FT      CDS
FT      Location/Qualifiers
        61..1014
        /tag="a
        /product="Human ApoE3"
XX
XX      JP2001017028-A.
XX
XX      PD
XX      23-JAN-2001.
XX
XX      28-APR-2000; 2000JP-0128919.
XX
XX      PR
XX      06-MAY-1999; 99JP-0125647.
XX
XX      PA
XX      (MITU ) MITSUBISHI CHEM CORP.
XX
XX      WI: 2001-285406/30.
XX      P-PSDB; AAB80997.
XX
XX      New apoe humanized mammalian cell useful for screening for agents
XX      useful for treating or preventing Alzheimer's disease and
XX      arteriosclerosis -
XX
XX      PS
XX      Disclosure; Page 13-14; 22pp; Japanese.
XX
XX      The present invention relates to an Apoe humanised mammalian cell. The
XX      present sequence is the coding sequence for human ApoE3, which was used
XX      in the method of the present invention. The Apoe humanised mammalian cell
XX      can be used for screening for agents useful for treating or preventing
XX      Alzheimer's disease and arteriosclerosis.
XX
XX      Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other:
XX
XX      Alignment Scores:
XX      Pred. No.: 4,64e-115 Length: 1156
XX      Score: 1577.00 Matches: 316
XX      Percent Similarity: 99.68% Conservative: 0
XX      Best Local Similarity: 99.68% Mismatches: 1
XX      Query Match: 99.24% Indels: 0
XX      DB: 22 Gaps: 0

```

```

US-09-827-854-16 (1-317) x AAF84315 (1-1156)
OY      1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
        |||
        61 ATGAAGGTTCTGTGGCGCTGGCTGGTGCACATTCTCTGGCAGAGATGCCAGGCCAAGGTG 120
OY      21 GluGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
        |||
        121 GAGCAAGCGGTGAGACAGAGCCGAGCTGCCGACAGACAGACGAGTGGCAGAGC 180
OY      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
        |||
        181 GGCACGCGCTGGGAACCTGGCACTGGGTGGCTTTGGATTACTCTGCTGGGTGGCAGACA 240
OY      61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
        |||
        241 CTGTGTGACAGAGTGCAGAGAGAGCTGCTCACTCCAGATCACCAGAACTGAGGGCG 300
OY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
        |||
        301 CTGATGACAGACCATGACATGAGAGTGAAGGCCCTCAATCGGAAGCTGGAGAACACTG 360
OY      101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
        |||
        361 ACCCGGTGGCGAGAGACAGCGGCGACGCTGTCCAAAGAGAGTCCAGAGCGCGCGAGGCC 420

```


QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 DB 421 CGGCTGGCGCGGACATGAGGACGTGTGGGGCCGCTGTGACGATACCGCGCGAGG 480
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 481 CAGGCCATGCTCGCGCAGAGACCGAGAGCTGGGGCTCCCTCCACCTGCGC 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 DB 541 AACCTGCGTAAAGGGCTCTCTCCGCGATGCCGATGACCTGCAGAAAGCGCTGGAGTAC 600
 QY 181 GlnAlaGlyAlaArgGlyGluGlyArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 601 CAGGCGGGGCGCGAGGCGCGGAGCGGCGCTCAGCGCCATCCGCGAGCGCTGGGG 660
 QY 201 ProLeuValGluGlnGlnLysArgValArgAlaIleThrValGlySerLeuAlaGlyGlnPro 220
 DB 661 CCCCTGGTGACAGGCGCGCTGCGGGCGCCACTGTGGGCTCCTGGCGGCGCAGCG 720
 QY 221 LeuGlnGluArgAlaGlnAlaThrArgGlyLeuArgLeuArgAlaArgMetGluMetGly 240
 DB 721 CTACAGAGAGCGGCGCCAGGCTGTGGCGAGCGGCTGCGCGCGGATGAGAGATGGCG 780
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 DB 781 AGCGGACCGCGAGCGCTGTGACGAGGTGAAGAGCAGAGTGCGGAGGTGGCGCCAG 840
 QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleArgLeuLys 280
 DB 841 CTGGAGAGCAGGCGCCAGAGATACCTGTGACGCGCGCTGTCCAGGCGCGCTCAG 900
 QY 281 SerTrpGluGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 DB 901 AGCTGGTTCGAGCGCTGTGGGAAAGCATGCACGCCAGTGGCGCGGCTGTGGAGAG 960
 QY 301 ValGlnAlaAlaValGlyThrSerAlaIleProValProSerAspAsnHis 317
 DB 961 GTGACAGGTGCGGTGGGACACAGCGCGCCCTGTGCGCCAGCAGCATATCAC 1011
 RESULT 8
 AAD22048
 ID AAD22048 standard; DNA; 1156 BP.
 AC AAD22048;
 DT 12-FEB-2002 (first entry)
 DE Human apolipoprotein E (apoE) isoprotein, apoE3 DNA.
 KW Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;
 KM hypertriglyceridaemia; low density lipoprotein; LDL; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 61..1014
 FT sig_peptide 61..114 /product= "Human apoE isoprotein, apoE3"
 FT mat_peptide 115..1011 /tag= b
 FT /tag= c
 FT /product= "Mature human apoE isoprotein, apoE3"
 PN WO200177136-A1.
 PD 18-OCT-2001.
 PF 06-APR-2001; 2001WO-US11358.
 PR 06-APR-2000; 2000US-0544386.

PR 04-OCT-2000; 2000US-0679088.
 PR 05-APR-2001; 2001US-0827854.
 PA (KOSP-) KOS PHARM INC.
 PA (UYBO-) UNIV BOSTON.
 PI Zannis VI, Kypreos KE;
 DR WPI: 2002-010885/01.
 DR P-PSDB; AAE13294.
 PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
 PT cholesterol, delaying the onset of or treating atherosclerosis in
 PT mammal, without inducing hypertriglyceridemia
 PS Claim 14; Page 81; 91pp; English.
 CC The present sequence is a human apolipoprotein E (apoE)
 CC isoprotein, apoE3 DNA. The apoE lipoproteins are useful for
 CC lowering cholesterol, delaying the onset of atherosclerosis,
 CC treating or regressing atherosclerosis without inducing
 CC hypertriglyceridaemia, in a mammal lacking an endogenous,
 CC normally functioning apoE gene or low density lipoprotein (LDL)
 CC receptor or is at risk for developing atherosclerosis due to
 CC accumulation of lipoprotein remnants in the bloodstream or having
 CC a defect in remnant removal.
 SQ Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other;
 Alignment Scores:
 Pred. No.: 4,64e-115 Length: 1156
 Score: 1577.00 Matches: 316
 Percent Similarity: 99.68% Conservative: 1
 Best Local Similarity: 99.68% Mismatches: 0
 Query Match: 99.24% Indels: 0
 DB: Gaps: 0
 US-09-827-854-16 (1-317) x AAD22048 (1-1156)
 QY 1 MetLysValLeuTrpAlaIleLeuValIleThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 61 ATGAAAGTTCGTGGGCTGCGTGTGTCACATTCCTGCGAGAGATGCCAGGCGAAGG 120
 QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnThrGluTrpGlnSer 40
 DB 121 GAGCAAGCGGTGGAGACAGAGCGCGGAGCCGAGCTGCGCAGACAGACAGTGGCGAG 180
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB 181 GGCAGAGCTGGGAACTGGCACTGGGTCCCTTTGGGATTACCTGGCTGGGTGCAGACA 240
 QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB 241 CTGTCTGACAGAGTGCAGAGAGAGCTGTCACTGCCAGGTACACCAGAACTGAGAGCG 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluLeu 100
 DB 301 CTGATGACGAGACCAATGAAGAGTGAAGGCTTACAAATCGGAATCGAGAAACAATG 360
 QY 101 ThrProValAlaGlnGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIleAla 120
 DB 361 ACCCGGTGGCGAGAGAGAGCGGGGACGCTGTCCAAAGAGCTGACGAGCGCGCGAGCC 420
 QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 DB 421 CGGCTGGCGCGGACATGAGGACGTGTGGGGCCGCTGTGACGATACCGCGCGAGG 480
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 481 CAGGCCATGCTCGCGCAGAGACCGAGAGCTGGGGCTCCCTCCACCTGCGC 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180

Db 541 AAGCTGCGTAAGCGGCTCCTCCGCGATGCCGATGACCTGCAGAACGCCCTGGCAGTGTAC 600
QY 181 GlnAlaGlyAlaArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 200
Db 601 CAGGCCGGGGCCCCGGAGGGGGCCCCGAGCCGCTCAGCGCATCCGCGAGCGCTGGGG 660
QY 201 ProLeuValGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 220
Db 661 CCCCTGTGTAACAGAGCG 720
QY 221 LeuGlu 240
Db 721 CTACAGGAGCGGGCG 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGluGluGluGluGluGluGluGluGlu 260
Db 781 ACCCGGACCCCG 840
QY 261 LeuGlu 280
Db 841 CTGAGAGGAGCAGCG 900
QY 281 SerTTPheGluProLeuValGluAspMetGlnArgGlnTTPAlaGluGluGluGluGlu 300
Db 901 AGCTGTTCGAGCCCTGTGTGAGAACATGACAGCCGACATGCGCGCGCGCGCGCGCGCG 960
QY 301 ValGlnAlaAlaValGluThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAAGGCTGCGGTGGGACACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1011

RESULT 9

AAD22052 standard; DNA: 1156 BP.

XX AC AAD22052;
XX 12-FEB-2002 (first entry)
XX Human apolipoprotein E (apoE) allele, apoE2** DNA.
XX Human apolipoprotein E (apoE) allele, apoE2** DNA.
XX Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;
XX hypertriglyceridaemia; low density lipoprotein; LDL; ds.
XX Homo sapiens.
XX Location/Qualifiers
FH CDS
FT 61..1014
FT /*tag= a
FT /product= "Human apoE allele, apoE2**"
FT sig_peptide 61..114
FT /*tag= b
FT mat_peptide 115..1011
FT /*tag= c
FT /product= "Mature human apoE allele, apoE2**"
PN WO20017136-A1.
PD 18-OCT-2001.
PF 06-APR-2001; 2001WO-US11358.
PR 06-APR-2000; 2000US-0544386.
PR 04-OCT-2000; 2000US-0679088.
PR 05-APR-2001; 2001US-0827854.
XX (KOSP-) KOS PHARM INC.
XX (UYBO-) UNIV BOSTON.
XX Zannis VI, Kyriacos KE;
XX WPL: 2002-010885/01.
XX P-PSDB: AAE13298.

PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT cholesterol, delaying the onset of or treating atherosclerosis in
PT mammal, without inducing hypertriglyceridaemia -
PS Claim 14; Page 83; 91pp: English.
XX
CC The present sequence is a human apolipoprotein E (apoE)
CC allele, apoE2** DNA. The apoE lipoproteins are useful for
CC lowering cholesterol, delaying the onset of atherosclerosis,
CC treating or regressing atherosclerosis without inducing
CC hypertriglyceridaemia, in a mammal lacking an endogenous,
CC normally functioning apoE gene or low density lipoprotein (LDL)
CC receptor or is at risk for developing atherosclerosis due to
CC accumulation of lipoprotein remnants in the bloodstream or having
CC a defect in remnant removal.
XX
SQ Sequence 1156 BP; 207 A; 369 C; 432 G; 148 T; 0 other;

Alignment Scores:

Pred. No.:	9,56e-115	Length:	1156
Score:	1573.00	Matches:	315
Percent Similarity:	99.68%	Conservative:	1
Best Local Similarity:	99.37%	Mismatches:	0
Query Match:	98.99%	Indels:	0
DB:	24	Gaps:	0

US-09-827-854-16 (1-317) x AAD22052 (1-1156)

QY 1 MetLysValLeuTTPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGGTCTGTGGGCTCGTTGCTGTCATATCTCCGAGATGCCAGGCCAAGGTG 120
QY 21 GluGlnAlaValGluThrLeuProGluProGluLeuArgGluGlnThrLeuTTPGlnSer 40
Db 121 GAGCAAGCGGTGGAGACAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 41 GlyGlnArgTTPGluLeuAlaLeuGluGluArgPheTTPAspTyrLeuArgTTPValGlnThr 60
Db 181 GGCAGCGCTGGGAATGCTGCTGCTTTGGGATTAATCTGCGCGGGTGGAGACA 240
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGTGTGACAGGTGACAGAGAGCGTGCACCTCCAGTCACCCAGGACTGAGAGCGG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu 100
Db 301 CTGATGGACGAGACCATGAAGAGTTGAAGGCTTAAATCGAATGAGGAGAACACTG 360
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 361 ACCCGGTGGCGGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 421 CGGCTGGCGCGGACATGAGAGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGTG 480
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTGGCGCAGACACCGAGAGCTGCGGGTGGCTGCTCCACCTGGCGC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 541 AAGCTGCGTACAGCGCTCCTCCGCAATGCCATGACCTCAGAGAGCGCTGGCAGTGTAC 600
QY 181 GlnAlaGlyAlaArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 200
Db 601 CAGGCCGGGGCCCCGGAGGGGGCCCCGAGCCGCTCAGCGCATCCGCGAGCGCTGGGG 660
QY 201 ProLeuValGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 220
Db 661 CCCCTGTGTAACAGAGCG 720
QY 221 LeuGlnGluArgAlaGlnAlaTTPGlyGluArgGluLeuArgAlaArgMetGluGluMetGly 240

QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 901 AGTGGTTGGAGCCCTGGTGGAGACATGCAAGCCGAGTGGCCGCGCTGGTGGAGAG 960
QY 301 ValGlnAlaValAlaValGlyThrSerAlaAlaProValProSerAspSnhIs 317
DB 961 GTGACAGCTGCGCTGGGACACAGCCGCCCTGTGTCCAGACATAC 1011
RESULT 11
AAT06957
ID AAT06957 standard: cDNA to mRNA: 1157 BP.
AC AAT06957;
XX
DT 19-JUN-1996 (first entry)
XX
DE Human apolipoprotein-E (APOE) cDNA.
XX
KW Recombinant; human; apolipoprotein-E; APOE; insect cells; larva;
KW Manduca sexta; Autographica californica nuclear polyhedrosis virus;
KW haemolymph; lipid complex; biologically active; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT MISC_feature 62..119
FT note="misc_signal"
FT 62..1015
FT /*tag= b
FT
XX US5472858-A.
XX
XX 05-DEC-1995.
XX
XX 04-JUN-1991: 9105-0709949.
XX
XX 04-JUN-1991: 9105-0709949.
XX
XX (WISC) WISCONSIN ALUMNI RES FOUND.
XX
XX Attie AD, Beckage NE, Gretsch DG, Sturley SL;
XX
XX MPI: 1996-029812/03.
XX
XX P-PSDB: AAR86791.
XX
XX Prodn. of recombinant apo:lipoprotein E in insects - by infecting
PT Manduca sexta larvae with recombinant Autographica californica
PT nuclear polyhedrosis baculovirus vector.
XX
XX
XX Disclosure: Columns 11-14: 10pp; English.
XX
XX Recombinant human apolipoprotein-E (APOE) (AAR86791) can be produced
CC by preparing a genetic construct (contg. an APOE-encoding sequence,
CC e.g. AAT06957, and flanking regulatory sequences enabling the protein
CC to be expressed in insect cells), which is then introduced into a
CC Manduca sexta larva (using a recombinant Autographica californica
CC nuclear polyhedrosis virus) and recovering the protein from the
CC haemolymph of the larval host. The APOE produced is in a form
CC sufficiently complexed with lipids to be biologically active, which
CC cannot be achieved in insect cell cultures, and can therefore be
CC used in therapeutic applications.
XX
XX
SQ Sequence 1157 BP: 212 A; 370 C; 426 G; 149 T; 0 other;

Alignment Scores:
Pred. No.: 2,366-114
Score: 1568.00
Percent Similarity: 99.05%
Best local Similarity: 99.05%
Query Match: 98.68%
DB: 17
Length: 1157
Matches: 314
Conservative: 0
Mismatches: 3
Indels: 0
Gaps: 0

US-09-827-854-16 (1-317) x AAT06957 (1-1157)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlySerGlnAlaLysVal 20
DB 62 ATGAAGTTCTGTGGGCTCGTGGTGGTGCATCTTCTGGCAGAGCCAGGCAAGGTG 121
QY 21 GlnGlnAlaValAlaGlnTrpGluProGluProGluLeuArgGlnGlnTrpGlnSer 40
DB 122 GAGCAAGCCGTTGAGAGACAGACAGCCGAGCCGAGCTCCGACAGACCCAGTGGCAGAGC 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlnArgPheTrpAspTrpLeuArgTrpValGlnTr 60
DB 182 GGCACAGCTGGAGAACTGGACACTGGGTGCTTTGGATTACTGCGTGGGTGGCAGACA 241
QY 61 LeuSerGlnGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 242 CTGTCTGACAGAGTGCAGAGAGAGCTGCAGCTCCCAAGTCAACCAAGAACTGAGAGGGG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnGlnLeu 100
DB 302 CTGATGAGACAGACCATGAAGAGATTGAAGCCTACAAATCGAAGCTGAGAGAACACTG 361
QY 101 ThrProValAlaGlnGluThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
DB 362 ACCCCGCTAGCGAGAGAGACGCGCGCAGCTGTCCAAAGAGCTGCAGAGCGCAGGCC 421
QY 121 ArgLeuGlnAlaAspMetGluAspValLysGlyArgLeuValGlnTrpArgGlyGluVal 140
DB 422 CGGCTGGGGCGGACATGGAGAGAGCTGCGCGCGCTGTGATGATACCGCGGCGAGGTG 481
QY 141 GlnAlaMetLeuGlnLysSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 482 CAGGCCATGCTGGCCAGAGACACCCAGAGAGCTGCGGTGCTGCCCTCCCACTGCGCC 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
DB 542 AAGCTGCTGAAGGGCTCTCCGCGCATCCGATGATGATGATGATGATGATGATGATGATG 601
QY 181 GlnAlaGlyAlaArgGlnGlnLysAlaGlnArgGlyLeuSerAlaLeuArgGlnGlnGly 200
DB 602 CAGGCGGGGCGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 661
QY 201 ProLeuValGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
DB 662 CCCCTGTGTGAACAGGCG 721
QY 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnMetGly 240
DB 722 CTACAGAGACCGGGCCGAGGCTGGGGCGAGCGGCTCGCGCGCGGATGAGAGATGGGC 781
QY 241 SerArgTrpArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaLys 260
DB 782 AGTGGACCCGCGACCGCTGGAGAGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 841
QY 261 LeuGlnGlnGlnAlaGlnGlnLysArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 842 CTGGAGAGACAGAGCCAGAGATAGCCCTGCAAGCGCGAGAGCGCTCCAGGCGCGCGCT 901
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 902 ACTGTGTTGACCCCTGCTGGAGACATGACAGCGCCCAATGGGCGCGCTGTGGAGAG 961
QY 301 ValGlnAlaValAlaValGlyThrSerAlaAlaProValProSerAspSnhIs 317
DB 962 GTGACAGCTGCGCTGGGACACAGCCGCCCTGTGTCCAGACATAC 1012
RESULT 12
ABN95746
ID ABN95746 standard: DNA: 1157 BP.
XX
AC ABN95746;
XX
DT 13-AUG-2002 (first entry)

XX DE Gene #2244 used to diagnose liver cancer.
XX KM Gene: liver cancer; ds: hepatocellular carcinoma; hepatotropic;
XX KM metastatic liver tumor; cytostatic; expression profile; disease state;
XX KM disease progression; drug toxicity; drug efficacy; drug metabolism.
XX OS Homo sapiens.
XX PN WO200229103-A2.
XX PD 11-APR-2002.
XX PF 02-OCT-2001; 2001WO-US30589.
XX PR 02-OCT-2000; 2000US-237054P.
XX PA (GENE-) GENE LOGIC INC.
XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX DR WPI; 2002-426119/45.
XX PT Diagnosing and detecting the progression of liver cancer,
XX PT hepatocellular carcinoma or metastatic liver tumor in a patient,
XX PT involves detecting the level of expression of two or more genes in a
XX PT liver tissue sample -
XX PS Claim 1: SEQ ID NO 2244; 298bp; English.
XX CC The invention relates to a novel method for diagnosing and detecting the
XX CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX CC tumor in a patient, and differentiating metastatic liver cancer from
XX CC hepatocellular carcinoma in a patient, involving detecting the level of
XX CC expression of two or more genes represented in ABN93503-ABN97455 in a
XX CC tissue sample. The method of the invention has hepatotropic, and
XX CC cytosolic activity. The method is useful for diagnosing and detecting
XX CC the progression of liver cancer, hepatocellular carcinoma and metastatic
XX CC liver carcinoma in a patient. The method is useful for identifying
XX CC expression profiles which serve as useful diagnostic markers as well as
XX CC markers that can be used to monitor disease states, disease progression,
XX CC drug toxicity, drug efficacy and drug metabolism.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pot_sequences.
XX SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;
Alignment Scores:
Pred. No.: 2.36e-114 Length: 1157
Score: 1568.00 Matches: 314
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 99.05% Mismatches: 3
Query Match: 98.68% Indels: 0
DB: 24 Gaps: 0
US-09-827-854-16 (1-317) x ABN95746 (1-1157)
QY 1 MetLysValLeuTrpAlaIleuLeuValThPheLeuAlaGlyCysGlnAlaIleVal 20
DB 62 ATAAAGGTTCTGGGGCTGGTGGCTGGTCAATCTCTGGCAGAGATGCCAGGCAAGGTG 121
QY 21 GUGLNAIAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 122 GAGCAAGCGGTGAGACAGAGCGCGAGCCGAGCTGGCCAGACAGACCGAGTGGCAGAGC 181
QY 41 GlyGlnArgTrpGluLeuAlaIleuGlyArgPheTrpAspTyrLeuArgTyrValGlnThr 60
DB 182 GGCACACGCTGGGACACTGGCACTGGCTTCCTTTGGGATTAACCTGGCGTGGTGACAGA 241
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 242 CTGTCTGAGCAGGTGACGAGGAGCTGCTCACTCCCAACTCACCACCAAGAACTGAGGCG 301

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
DB 302 CTGATGAGCAGACACCATGAAGAGAGTTGAAGCGCTACAAATCGGAACCTGGAGGAACAACCTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIleGlnAla 120
DB 362 ACCCGGTACCGGAGAGACCGGGCCAGCGCTCTCCCAAGAGGCTGCAGACGGCGCAGGGCC 421
QY 121 ArgLeuGluValAspMetGluAspValCysGlyArgGluValGlnTyrArgGlyGluVal 140
DB 422 CGGCTGGGGCCGACATGAGAGACGTGTGGCCCGCTGTGACATGCCCGCGAGGTG 481
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 482 CAGGCGCATGCTGGCCACAGACACCGAGAGACTGTGGGTGGCGCTCCGCTCCACCTGGCGC 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
DB 542 AAGCTGGCTAAGCGGCTCTCCCGATCCCGATGACCTGCAGAAAGCCCTGCGCAGTGTAC 601
QY 181 GlnAlaGluValaArgGluGluGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 602 CAGGCGGGGCGCGAGGCGCGAGCGCGCTCAGCGCATCCGCGAGCGCTGGGG 661
QY 201 ProLeuValGluGlnGluArgValArgAlaAlaThrValGlySerLeuAlaGluGlnPro 220
DB 662 CCCCTGGTGACAGAGCGCGCGCTGGCGCGCCACTGTGGCTCCCTGGCGCGCGCAGCGC 721
QY 221 LeuGlnGluArgAlaGlnAlaIleTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
DB 722 CTACAGAGAGCGGGCCCGAGCGCTGGGGCGACCGCTGGCGGATGAGAGAGTGGCG 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValaIleArgValaArgAlaLys 260
DB 782 AGTGGACCGCGCACCGCTGGACGAGGTGAAGAGCAGGTGGCGAGGTGGCGCCAG 841
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
DB 842 CTGAGAGAGCAGCGCCAGCAGATAGCGCTCAGCGCGAGCGCTTCAGGCGCGCGCTCAAG 901
QY 281 SerTrpPheGluProLeuValaGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 902 AGCTGTGTTCAGAGCCCTCGTGGAAGACATGCACCGCAGTGGCGCGGTGTGGAGAG 961
QY 301 ValGlnAlaIleValaGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 962 GTGCAGGCTGCGCTGGGACACAGCGCGCGCTGTGTGCCAGCAATCAAC 1012
RESULT 13
ABK64514
ID ABK64514 standard; DNA; 1157 BP.
XX AC ABK64514;
XX 18-JUN-2002 (first entry)
XX DE Human benign prostatic hyperplasia gene #409.
XX KW Human: benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
XX OS Homo sapiens.
XX PN WO200212440-A2.
XX PD 14-FEB-2002.
XX PF 07-AUG-2001; 2001WO-US24708.
XX PR 07-AUG-2000; 2000US-223323P.
XX PR 05-JUN-2001; 2001US-0873319.
XX PA (GENE-) GENE LOGIC INC.


```
XX DE Sequence encoding human apolipoprotein E.
XX XX Hyperlipaemia; ds.
XX KM
XX OS Homo sapiens.
XX FH
XX Key Location/Qualifiers
XX CDS 15..968
XX FT /*tag= a
XX FT 69..965
XX FT mat_peptide /*tag= b
XX FT
XX JP60118189-A.
XX PN
XX XX
XX PD 25-JUN-1985.
XX PF 29-NOV-1983; 83JP-0224980.
XX PR 29-NOV-1983; 83JP-0224980.
XX PA (MITU ) MITSUBISHI CHEM IND KK.
XX DR WPI; 1985-188003/31.
XX DR P-PSDB; AAP51204.
XX XX
XX PT DNA fragment - contg. DNA which codes human apolipoprotein E for
XX treatment of hyperlipaemia.
XX PS Claim 3; Page 484; 8pp; Japanese.
XX CC The sequence may be used to produce the apolipoprotein E, useful in the
XX treatment of hyperlipaemia.
XX SQ Sequence 1110 BP; 198 A; 355 C; 414 G; 143 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 2.7e-114
Score: 1567.00 Length: 1110
Percent Similarity: 99.05% Matches: 314
Best Local Similarity: 99.05% Mismatches: 0
Query Match: 98.62% Indels: 3
DB: 6 Gaps: 0
US-09-827-854-16 (1-317) x AAN50450 (1-1110)
OY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIleVal 20
DB 15 ATGAAGTTCTGTGGGCTCGTTGCTGGTCACATTCCTGGCAGAGATGCCAGGCCAAGTGT 74
OY 21 GluGlnAlaIleValGluThrGluProGluLeuAlaGlnGlnThrGluTrpGlnSer 40
DB 75 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGTCCGACAGACAGACCAGTGGCAGAGC 134
OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 135 GGCACAGCGGTGGAGATGGCACTGGTCCCTTTTGGGATTACCTGCGCTGGGTGCAGACA 194
OY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 195 CTGTCTGAGACACGTGCAGAGAGCTGCTCACGTCCAGAGTCAACCAAGAACTGAGGGCG 254
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
DB 255 CTGATGGAGACACCATGAAAGAGTTGAAGGCTTACAAATCGAACTGGAGAGAACACTG 314
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIleGlnAla 120
DB 315 ACCCGGTGGCGGAGAGAGCGGGCGGCTGTCCAAAGAGACTGCAGGCGGCCAGGCC 374
OY 121 ArgLeuGlyAlaAspMetCysIleValCysGlyArgLeuValGlnTyrArgGlyVal 140
DB 375 CGGCTGGGCGCGACATGAGAGACGTGTCCGCGCTGTGTCAGTACCGCGGCGAGGTG 434
```

```
OY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 435 CAGGCCATCTGGGCCAGAGACACCGAGAGCTGGGGGTGCGCTCCCTCCACCTGCCGC 494
OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
DB 495 AAGCTGCGTAAGCGGCTCTCCCGCATGTGACCTCCAGAAAGGCGCTGCAGAGTAC 554
OY 181 GlnAlaGlyAlaArgGluGluValArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 555 CAGGCCGGGGCCCGGAGGCGCCGAGCGGCGCTCCAGCCATCCGAGGCGCTGGGG 614
OY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 615 CCCCTGTGTGAACAGAGGCGCGTGGGCGCCGCTGTGGCTCCCTGSCCGGCCAGCCG 674
OY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
DB 675 CTACAGAGAGCGGGCCAGGCGCTGGGGCGAGCGGCTGGCGGATGAGAGATGGGC 734
OY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 735 ACCCGGACCCCGGACCGCTGGACGAGGTGAAGAGAGCAAGTGGCGAGGTGCGCGCCAAG 794
OY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 795 GTGAGAGAGACAGGCGCCAGATACGCTGCAGGCGGAGGCGCTTCCAGGCCCGCTCAAG 854
OY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 855 AGCTGTGTGACCCCTGTGTGAAGACATGACAGCGCCGACATGGGCGGCTGTGTGAGAA 914
OY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 915 GTCAAGCTGCGGTGGGACACAGCGCGCCCTGTGTCCAGGACATATCAC 965
```

Search completed: March 14, 2003, 12:43:10
Job time : 183.728 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:17:52 ; Search time 34.7343 Seconds
(without alignments)
2798.866 Million cell updates/sec

Title: US-09-827-854-16

Perfect score: 1589

Sequence: 1 MKVLMALVFLAGCQAKV.....VEKQAAVGTSAAPVSDNH 317

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame.p2n.model -DEV=xl
-Q=/cgn2_1/USPTO_spool/US09827854/rnmat_11032003_101611_27510/app_query.fasta.1.3576
-DB=Issued_Patents.NA -QFWT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854 -RCGN_1_1_224 &rnmat_11032003_101611_27510 -NCRU=6 -ICRU=3
-NO_XLPTX -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PCRTUS.COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1568	98.7	1157	1	US-07-709-949-1
2	1130	71.1	1126	4	US-08-949-155-5
3	1130	71.1	1126	4	US-09-819-964-5
4	994	62.6	4267	4	US-08-949-155-51
5	994	62.6	4267	4	US-09-819-964-51
6	976	61.4	660	2	US-07-726-306A-28
7	503	31.7	330	1	US-07-849-389-6
8	374	23.5	252	3	US-08-617-256-24
9	374	23.5	252	4	US-09-287-141-24
10	374	23.5	252	4	US-09-431-613-24
11	374	23.5	252	4	US-09-504-245-24
12	374	23.5	252	4	US-09-287-682-24

13	374	23.5	252	4	US-09-287-679-24
14	374	23.5	252	4	US-09-397-768-24
15	374	23.5	252	4	US-09-287-681-24
16	374	23.5	252	4	US-09-495-444-24
17	176.5	11.1	842	1	US-08-952-736-1
18	173.5	10.9	801	1	US-07-959-946-4
19	173.5	10.9	801	1	US-08-333-577-4
20	173.5	10.9	801	5	PCT-US92-08634-4
21	159	10.0	5661	4	US-08-938-105-2
22	156.5	9.8	964	1	US-08-448-606-5
23	153.5	9.7	863	1	US-08-448-606-7
24	152.5	9.6	4296	4	US-09-060-410-3
25	150	9.4	3256	2	US-08-968-751-3
26	149	9.4	1879	4	US-09-750-580-2
27	146.5	9.2	603	4	US-08-952-796-14
28	146.5	9.2	4852	1	US-07-853-913-3
29	144	9.1	13121	4	US-08-961-527-126
30	141	8.9	6306	1	US-08-466-390-3
31	141	8.9	6306	1	US-08-470-950-3
32	141	8.9	6306	1	US-08-467-781-3
33	141	8.9	6306	1	US-08-195-487-3
34	141	8.9	6306	2	US-08-483-924-3
35	141	8.9	6306	5	PCT-US93-06160-3
36	141	8.9	8789	1	US-08-328-254-5
37	141	8.9	10136	5	PCT-US95-16216-2
38	141	8.9	10136	5	PCT-US95-16216-2
39	140	8.8	1771	2	US-08-533-669A-7
40	140	8.8	1771	2	US-08-511-872-1
41	140	8.8	1771	4	US-09-183-861-7
42	140	8.8	1771	4	US-09-022-765-7
43	140	8.8	1960	2	US-08-533-306A-1
44	140	8.8	1960	2	US-08-742-923A-1
45	140	8.8	4868	1	US-08-139-937-12

ALIGNMENTS

RESULT 1
US-07-709-949-1
; Sequence 1, Application US/07709949
; Patent No. 5472858
; GENERAL INFORMATION:
; APPLICANT: Attie, Alan D
; APPLICANT: Gretch, Daniel G
; APPLICANT: Sturley, Stephen L
; APPLICANT: Beckage, Nancy E
; TITLE OF INVENTION: Production of Recombinant Proteins in
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: P.O. Box 2113
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/709,949
; FILING DATE: 19910604
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J, 386
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 9629691801
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 1:

Sequence 24, Appl
Sequence 24, Appl
Sequence 24, Appl
Sequence 24, Appl
Sequence 1, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 2, Appl
Sequence 5, Appl
Sequence 7, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 2, Appl
Sequence 14, Appl
Sequence 3, Appl
Sequence 126, App
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 7, Appl
Sequence 1, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 1, Appl
Sequence 12, Appl

SEQUENCE CHARACTERISTICS:
LENGTH: 1157 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 62..1015
FEATURE:
NAME/KEY: misc_signal
LOCATION: 62..119
PUBLICATION INFORMATION:
AUTHORS: McLean, J W
JOURNAL: Journal of Biological Chemistry
VOLUME: 259
PAGES: 6498-6504
DATE: 1984
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 1157
US-07-709-949-1

Alignment Scores:
Pred. No.: 4,54e-144 Length: 1157
Score: 1568.00 Matches: 314
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 99.05% Mismatches: 3
Query Match: 98.68% Indels: 0
DB: 1 Gaps: 0

US-09-827-854-16 (1-317) x US-07-709-949-1 (1-1157)

QY 1 MetLysValLeuTrpAlaAlaLeuValThPheLeuAlaGlyCysGlnAlaLysVal 20
DB 62 ATGAGGTTCTGTGGCTGTGGCTGTGCATCTCTGCGAGAGTCCAGAGGCTG 121

QY 21 GluGlnAlaValAlaGluTrpGluProGluLeuArgGlnGlnTrpGlnIns 40
DB 122 GACAGAGCGTGGAGAGAGCGGAGCCGAGCTGGCGAGAGCGAGCGAGTGGAG 181

QY 41 GlyGlnArgTrpAlaLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB 182 GCGCAGCGCTGGAGAGTGGAGTGGCTTTGGGATTACTGGCGTGGTGGAGACA 241

QY 61 LeuSerGluGlnValGlnGluGluLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 242 CTGTCTGAGCAGGTGAGGAGGAGCTGCTCAGCTCCCAAGTCCCAAGAACTGAGGGCG 301

QY 81 LeuMetAspGluTrpMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnIns 100
DB 302 CTGATGAGCAGACCATGAGAGGTTGAGAGCCTACAAATCGAAGTGGAGAGACTG 361

QY 101 ThrProValAlaGlnGluTrpArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
DB 362 ACCCGGTAGCGAGAGAGAGCGGGGACGGCTGCCAAGAGCTGCAGAGCGGGAGGCC 421

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
DB 422 CGCGTGGCGCGGACATGAGGAGCTGTGGCGCGCTGTGCAGTACCCGCGGAGGTG 481

QY 141 GlnAlaMetLeuGlnInsTrpGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 482 CAGGCCATGCTGGCCAGAGACACCGAGGAGCTGGGTGGCTCCCTCCCTCCAGTGGCG 541

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
DB 542 AAGCTCGTAAGAGGGGCTCCCGCGCATCCGATCCGAGACTGCAGAAAGCGGCTGGAGTAC 601

QY 181 GlnAlaGlyAlaArgGluGlnGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200

DB 602 CAGCGCGGGCCCGCGAGGGCGGAGCGGCGCTCAGCCCATCCGCGAGCCCTGGGG 661

QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220

DB 662 CCCCTGTGTAACAGAGGCGCGCTGGCGGCGCCACTGTGGGCTCCCGGCGGCGAGCG 721

QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGlyArgLeuArgAlaArgMetGluMetGly 240

DB 722 CTACAGAGGCGGCGCGCTGGCGGAGCGGCTGCGCGCTGCGAGTGGAGAGTGGCG 781

QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260

DB 782 AGTGGAGCCCGCAGCCGCTGGAGAGGTGAGAGAGAGGTGGGAGGCGCGCCAG 841

QY 261 LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnGlnAlaGluValPheGlnAlaArgLeuLys 280

DB 842 CTGAGAGAGCAGGCGCCAGCAGATACGCTGACAGCGAGCGCTCCAGGCGCGCTCAAG 901

QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGlyLys 300

DB 902 AGCTGTTCCAGCCCTGCTGGGAGACATCCAGCGCAGTGGCGCGGCTGTGGAGAG 961

QY 301 ValGlnAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317

DB 962 GTGAGGCTGCGTGGGACACAGCGCGCCCTGTGCGCAGCACAATCAC 1012

RESULT 2
US-08-949-155-5
Sequence 5, Application US/08949155
Patent No. 6271436
GENERAL INFORMATION:
APPLICANT: Piedrahita, Jorge A
APPLICANT: Bazer, Fuller W
TITLE OF INVENTION: Compositions and Methods for the
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESS: ARNOLD, WHITE AND DURKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/949,155
FILING DATE: Concurrently Herewith
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/027,338
FILING DATE: 11-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/046,094
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hibler, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TAAK:177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:


```

NAME/KEY: CDS
LOCATION: 51..1001
US-08-949-155-5

Alignment Scores:
Pred. No.: 2,18e-101
Score: 1130.00
Percent Similarity: 83.44%
Best Local Similarity: 69.69%
Query Match: 71.11%
DB: 4
Gaps: 3

US-09-827-854-16 (1-317) x US-08-949-155-5 (1-1126)

OY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 51 ATGAGGGTTCTGTGGTCTTGTGGTAACTCTCTCGCAGATGCGGACAGAGAC 110
OY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
DB 111 GAGCCGGGG-----CCGCGCGGAGGTGCACGTGTGTGGAGAGAGCCCAAG 158
OY 38 TrpGlnSerGlyGlnAlaTrpGluLeuAlaLeuGlyAtrPheTrpAspTrpLeuArgTrp 57
DB 159 TGGCAGGGCAGCAGCCCTGGGAGCAGCCCTGCGGCTCTGCGATTACCTGCGCTGG 218
OY 58 ValGlnThrLeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGlu 77
DB 219 GTGAGTCCCTGTCTGACCAAGTCAGAGAGAGTGTCTACACCAAGTCAAGCCAGGAA 278
OY 78 LeuAlaAlaLeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlu 97
DB 279 CTGACGAGAGCTGATGAGAGAGAGCATGAAGAGGTGAAGCCCTACCGGAGAGAGCTGAG 338
OY 98 GlnGlnLeuThrProValAlaGluGluThrArgAlaAtrGluSerLysGluLeuGlnAla 117
DB 339 GCGCAGCTGGGGCCCGTACCAGCAGAGCAGCGCCCTCTCCAAAGACTGCGAGCGC 398
OY 118 AlaGlnAlaArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArg 137
DB 399 GCGCAGGCGCCGCTGGGCGCGACATGAGAGCGTGCACCAACGCTTGTCTCTACCGC 458
OY 138 GlyGlnValGlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSer 157
DB 459 AGCGAGGTGCACAACATGTTGGGCGCAGACACCGAGAGCGTGGAGCGCTGCTTCC 518
OY 158 HisLeuAtrGlyLeuAtrLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeu 177
DB 519 CACCTGCGCAAGCTGCGCAGCGGCTGCTCCGACACCGAGAGACTGCAGAAAGCGCTG 578
OY 178 AlaValTrpGlnAlaGluAlaArgGlnGluArgGlyLeuSerPheAlaIleArgGlu 197
DB 579 GCGGTGTACCAAGCGGGGCTGCGCAGGCGCGCAGCGCTGAGGCGCTCCGCTCAG 638
OY 198 ArgLeuGlyProLeuValGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAla 217
DB 639 CGCCTCGGGGCGCTGTGTGAGAGAGCGCGATTTGCGCGCGCCACCTAGTACCAAGGCGC 698
OY 218 GlyGlnProLeuGlnGluArgAlaGlnAlaAtrPglGlyLeuArgLeuArgAlaArgMetGlu 237
DB 699 GCGCAGCGCTGCGCAGCGCGCGGAGCTGCGGCGCGAGAGCTGCGGAGCGCTGAG 758
OY 238 GluMetGlySerArgTrpArgAspArgLeuAspGluValLysGlnGluValAlaGlnVal 257
DB 759 GAGATGGGACAGCGGACCGCGGACCGCTGATGATGATGATGATGATGATGATGATGATG 818
OY 258 ArgAlaLysLeuGlnGluGlnAlaGlnGlnLysLeuArgLeuGlnAlaAlaPheGlnAla 277
DB 819 CGCACCAGAGTGAAGAGAGAGAGCGCAGATTGGCGCTGCGAGCGCGAGAGATTCCAGCC 878
OY 278 ArgLeuLysSerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeu 297
DB 879 CTCCTCAAGGCTGTCTGAGCTCTCTGTGTGAGACATATACGCGCGAGAGAGAGAGAGAG 938

```

```

OY 298 ValGlnLysValGlnAlaAlaValGly---ThrSerAlaAlaProValProSerAspAsn 316
DB 939 GTGAGAGAGATGACGTGCGCGGTGAGATATAGCTCTCCACCTTGCAGGATGATAT 998

RESULT 3
US-09-819-964-5
Sequence 5, Application US/09819964
Patent No. 6369294
GENERAL INFORMATION:
APPLICANT: Piedrahita, Jorge A
Bazer, Fuller W
TITLE OF INVENTION: Compositions and Methods for the
Generation of Transgenic Animal Species
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESS: ARNOLD, WHITE AND DURKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/819,964
FILING DATE: 28-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/949,155
FILING DATE: <Unknown>
APPLICATION NUMBER: US 60/046,094
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hibler, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TMK:177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 51..1001
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-819-964-5

Alignment Scores:
Pred. No.: 2,18e-101
Score: 1130.00
Percent Similarity: 83.44%
Best Local Similarity: 69.69%
Query Match: 71.11%
DB: 4
Gaps: 3

US-09-827-854-16 (1-317) x US-09-819-964-5 (1-1126)

OY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 51 ATGAGGGTTCTGTGGTCTTGTGGTAACTCTCTCGCAGATGCGGACAGAGAC 110
OY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
DB 111 GAGCCGGGG-----CCGCGCGGAGGTGCACGTGTGTGGAGAGAGAGAGAGAGAGAG 158

```



```

QY 38 TrrpGlnSerGlylnArgrtrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrrp 57
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 159 TGGCAGGAGGAGCCAGCCCTGGGCGAGGCCCTGGCGCTTCGGGATTACCGCCCTGG 218
QY 58 ValGlnThrLeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGln 77
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 219 GTTCAGTCCCTGTCTACCAAGTCGAGAGAGAGCTGTCTACGACCAAGTCACCCAGGA 278
QY 78 LeuArgAlaLeuMetAspGlnThrMetLysGluLeuLysAlaTyrLysSerGluLeuGln 97
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 279 CTGACGAGCTGTATGAGGAGACATGAAAGAGAGTGAAGCCCTACCGCGAGCTGGAG 338
QY 98 GlnGlnLeuThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGluLeuGlnAla 117
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 339 GCGCAGGCGCCGCTGGGCGCCGACATGAGAGAGCTGCCAACCGCTTGGTCTCTACCGC 398
QY 118 AlaGlnAlaArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArg 137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 399 GCGCAGGCGCCGCTGGGCGCCGACATGAGAGAGCTGCCAACCGCTTGGTCTCTACCGC 458
QY 138 GlyGlnValGlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSer 157
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 459 AGCGAGGTGCACAAATGTTGGGCGAGACACCGAGAGCTGGGAGCGCGCTGGCTTCC 518
QY 158 HisLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeu 177
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 519 CACCTCGCAGAGCTGCGCAAGCGCGCTGCTCCGCGACACCGACCTGCGAGAACGCCCTG 578
QY 178 AlaValTyrGlnAlaGlnAlaArgGlnGlnGlnAlaArgGlyLeuSerAlaIleArgGln 197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 579 GCGCTTACAGGCGGCGCTGCGCGAGGCGCGGAGCGCGACCGCTGAGCGCCCTCCCGCAG 638
QY 198 ArgLeuGlyProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAla 217
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 639 CGCGCTGGGCGCCCTGTGTGAGAGAGCGCGCATTCGCGCGCGCCCTGAGTACAGAGGCC 698
QY 218 GlyGlnProLeuGlnGlnArgAlaGlnAlaTrrpGlyGlnArgLeuArgAlaArgMetGln 237
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 699 GCGCAGCCCTGCGCGAGCGCGCGGAGGCGCGGAGCGCTGGGCGCACAAAGCTGGCGAG 758
QY 238 GluMetGlySerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnVal 257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 759 GAGATGGGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 818
QY 258 ArgAlaLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 277
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 819 CCGACCAAGTGGAGGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 878
QY 278 ArgLeuLysSerThrPheGlnProLeuValGluAspMetGlnArgGlnThrAlaGlyLeu 297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 879 CTCCTCAAAGGCTGTTCAGAGCTGTGTGGAAGACATACGGCGCGCGAGCGCGCGCGCTG 938
QY 298 ValGlnLysValGlnAlaValGly--ThrSerAlaAlaProValProSerAspAsn 316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 939 GTGGAGAGGATGCAGTGGCGCGGTGAGCATACGCTCTCCACCTCTGCGCGCGAGTAAAT 998

```

RESULT 4
US-08-949-155-51
Sequence 51, Application US/08949155
Patent No. 6271436

GENERAL INFORMATION:

APPLICANT: Piedrahita, Jorge A
APPLICANT: Bazer, Fuller W
TITLE OF INVENTION: Compositions and Methods for the
TITLE OF INVENTION: Generation of Transgenic Animal Species
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE AND DURKEE
STREET: P. O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/949,155
FILING DATE: Concurrently herewith
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/027,338
FILING DATE: 11-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/046,094
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hilder, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TANK:177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ. ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 4267 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-949-155-51

```

```

Alignment Scores:
Pred. No.: 2,39e-87 Length: 4267
Score: 994.00 Matches: 217
Percent Similarity: 58.20% Conservative: 42
Best Local Similarity: 48.76% Mismatches: 47
Query Match: 62.56% Indels: 140
DB: 4 Gaps: 5

```

US-09-827-854-16 (1-317) x US-08-949-155-51 (1-4267)

```

QY 1 MetLysValLeuThrPalaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    ||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2448 ATGGCTGTGAATGCTGCTGCT-----CGAGAGAGCGCGAGACAGAGAGC 2489
QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArg-----GlnGlnThrGln 37
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2490 GAGCCGCGG-----CCGCGCGCGAGAGTGCACGTGTGTGGAGGAGGAGCCCAAG 2537
QY 38 TrrpGlnSerGlylnArgrtrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrrp 57
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2538 TGGCAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGG 2597
QY 58 ValGlnThrLeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGln 77
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2598 GTTCAGTCCCTGTCTGACCAAGTCGAGAGAGAGCTGTCTACGACCAAGTCACCCAGGA 2657
QY 78 Leu----- 78
Db 2658 CT-GACGTAAGTCCACCGACCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCTGA 2716
QY 78 ----- 78
Db 2717 CCCTCTGGCAACCGTGTCTGTGAGACCTCAGGCTCCACCGCGGTTCTCTCTG 2776
QY 78 ----- 78
Db 2777 TCCTTGTCCCAACTCTTGGGGGTCTGGGTCTCTGTTCTTTTCTCTCTCTCTTT 2836
QY 78 ----- 78
Db 2837 TTGGGGGAAAAAATTCTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCC 2896
QY 78 ----- 78

```



```

Db 2897 ATCTTGAGCTCCTGCTTCCCTGCTCTGAGTCTTCCCTGCTCTCTCTG 2956
Oy 78 ----- 78
Db 2957 AATCTCTGGACGCTCTGGCCATCCGACCTTACAGACCTCTCTCTCTCTG 3016
Oy 79 ----- ArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaThr 92
Db 3017 CCCGCCCTCTCTGCGCCAGGAGCTGATAGAGAGACATGAGAGAGTGAAGCCCTAC 3076
Oy 93 LysSerGluLeuGluGluGluLeuThrProValAlaGluGluThrArgAlaArgLeuSer 112
Db 3077 CCGGAGGAGCTGGAGCGCCAGCTGGCCCGCTGACCCAGAGAGACGACGCGCTGTCC 3136
Oy 113 LysGluLeuGlnAlaAlaGlnAlaArgLeuGlnAlaAspMetGluAspValCysGlyArg 132
Db 3137 AAGGAGCTGCAGGCGCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3196
Oy 133 LeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThrGluGluLeuArg 152
Db 3197 TTGGTGGCTTACCGGACGAGGAGTGCAACAATGTTGGGCGACACCGAGAGAGCTGCGG 3256
Oy 153 ValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAsp 172
Db 3257 AGCCGCTGGCTTCCCTGCGGCAAGCTGCGCAAGCGGCTGCTCCGCGACCGAGGAC 3316
Oy 173 LeuGlnLysCysLeuAlaValTyrGlnAlaGlnAlaArgGluGluValGluArgLysLeu 192
Db 3317 CTGCGAAGAGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3376
Oy 193 SerAlaIleArgGluArgLeuGlyProLeuValGluGlnGlyArgValArgAlaAlaThr 212
Db 3377 AGCGGCTTCCCGGACGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3436
Oy 213 ValGlySerLeuAlaGlyGlnProLeuGlnGluArgAlaGlnAlaTyrGlyGluArgLeu 232
Db 3437 CTGATCTACAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3496
Oy 233 ArgAlaArgMetGluGluMetGlySerArgThrArgAspArgLeuAspGluValLysGlu 252
Db 3497 CGCGGAGCGCTGGAGAGATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3556
Oy 253 GlnValAlaGluValArgAlaLysLeuGlnGluGlnAlaGlnIleArgLeuGlnAla 272
Db 3557 CAGCTGAGGAGAGCTGCGACCAACTGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3616
Oy 273 GluAlaPheGlnAlaArgLeuLysSerTyrPheGluProLeuValGluAspMetGlnArg 292
Db 3617 GAGGGAATTCACGCGCTCTCAAGCTGTGTCGAGCTCTGTGTGAAGACATACGCGGC 3676
Oy 293 GlnTyrPalaGlyLeuValGluLysValGlnAlaAlaValGly---ThrSerAlaAlaPro 311
Db 3677 CAGTGGGCGGCGGCTGGTGGAGAGATGAGTGCAGCTGTGAGCATTAAGCTCTCCACCTCT 3736
Oy 312 ValProSerAspAsn 316
Db 3737 GCGCCCACTGATTAAT 3751

```

```

STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/819,964
FILING DATE: 28-Mar-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/949,155
FILING DATE: <unknown>
APPLICATION NUMBER: US 60/046,094
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hibler, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TANK:177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 4267 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 51:
US-09-819-964-51
Alignment Scores:
Pred. No.: 2.39e-87 Length: 4267
Score: 994.00 Matches: 217
Percent Similarity: 58.20% Conservative: 42
Best Local Similarity: 48.76% Mismatches: 47
Query Match: 62.56% Indels: 140
DB: Gaps: 5
US-09-827-854-16 (1-317) x US-09-819-964-51 (1-4267)
Oy 1 MetLysValLeuThrPalaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 2448 ATGCTGTGAATCTGGCTGGT-----GCAGATGCCGCGACAGAGAC 2489
Oy 21 GlnGlnAlaValGluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
Db 2490 GAGCGGGG-----CCGCGCGCGGAGGTGTCAGCTGTGGTGGAGAGAGCCCAAG 2537
Oy 38 TyrGlnSerGlyGlnArgTyrPheGluLeuAlaGlnGlyArgPheTyrAspTyrLeuArgTyr 57
Db 2538 TGGCAGCGAGCAGCCCTCGTGGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2597
Oy 58 ValGlnThrLeuSerGluGlnValGlnGluLeuLeuLeuSerGlnValThrGlnGlu 77
Db 2598 GTGAGTCCCTGTGTGACCAAGTGAAGAGAGCTGCTCAGCACCAGGTACCCAGGAA 2657
Oy 78 Leu----- 78
Db 2658 CT-GACGTAAAGTCCCGACCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCTGA 2716
Oy 78 ----- 78
Db 2717 CCTCTGCGGAGACGTTGTGTGAGACCTCAGGCTCCACCCGTCGCGGTTCTCTGTG 2776
Oy 78 ----- 78
Db 2777 TCTTGTGCGCACTCTTGGGGGTGTGGGTCTGTCTTTCTTTCTTTCTTCTCTCTTTT 2836
Oy 78 ----- 78

```



```

Db 2837 TTGGGGGGAAAAAATTCTTTCTTTCTTTCAATTGACTTCATGCTTGTCTTCTTCC 2896
Oy 78 ----- 78
Db 2897 ATCTTGAGCTCTGCTGCGCTGTCTGTGGTCACTTGTGCGCTTGTCTGTCTG 2956
Oy 78 ----- 78
Db 2957 AATCTGTGCACTGCTGCGCATGCGACGTGAGAGCCCTCTTGTCCCTCACCGCC 3016
Oy 79 ----- 79
Db 3017 CCCGCCCTCTTCCGCCAGGAGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3076
Oy 93 LysSerGluLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 112
Db 3077 CCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3136
Oy 113 LysGluLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 132
Db 3137 AAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3196
Oy 133 LeuValGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 152
Db 3197 TTGCTCTCTACCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3256
Oy 153 ValArgLeuAlaSerHisLeuArgLysLeuArgLysLeuArgLysLeuArgLysLeu 172
Db 3257 AGCGGCTGCTGCTCCACCTCCGCAAGCTGCGCAAGCGCTGCTCCGCGACAGCGAGAG 3316
Oy 173 LeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 192
Db 3317 CTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3376
Oy 193 SerAlaIleArgGluArgLysLeuGluGluGluGluGluGluGluGluGluGluGlu 212
Db 3377 AGGCGCTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3436
Oy 213 ValGlySerLeuAlaGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 232
Db 3437 CTGAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3496
Oy 233 ArgAlaIleArgMetGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 252
Db 3497 CCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3556
Oy 253 GluValAlaGluValArgLysLeuGluGluGluGluGluGluGluGluGluGluGlu 272
Db 3557 CAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3616
Oy 273 GluAlaPheGluAlaArgLysSerTrpPheGluProLeuValGluAspMetGlnArg 292
Db 3617 GAGGAGTTCACAGCGCTCTCAAGGCTGTTGAGACCTCTGTGGTGAACATACGCGCC 3676
Oy 293 GlnTrpAlaGlyLeuValGluGluGluGluGluGluGluGluGluGluGluGluGlu 311
Db 3677 CAGTGGGCGGCGGTGGTGAAGAGATGATGCGCGGTGAGATAGATCTCTCCACTCT 3736
Oy 312 ValProSerAspAsn 316
Db 3737 GCGCCAGTGTAT 3751

```

```

ADDRESSER: Banner & Wilcoff, Ltd.
STREET: 1 Financial Center
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726.306A
FILING DATE: 02-Oct-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 95/20080.4
FILING DATE: 02-Oct-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/009.832
FILING DATE: 01-Jan-1996
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Ph.D., Kathleen M.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 345-9100
TELEFAX: (617) 345-9111
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-726-306A-28
Alignment Scores:
Pred. No.: 1,096-86 Length: 660
Score: 976.00 Matches: 196
Percent Similarity: 98.49% Conservative: 0
Best Local Similarity: 98.49% Mismatches: 3
Query Match: 61.42% Indels: 0
Gaps: 0
US-09-827-854-16 (1-317) x US-08-726-306A-28 (1-660)
Oy 1 MetLysValLeuTrpAlaIleLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAGAGTTCTGTGGGCTGCTGTTGCTGTCACATTCCTGCGAGAGTCCAGGCCAAGGTG 121
Oy 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGluGlnThrGluTrpGlnSer 40
Db 122 GAGCAAGCGGTGAGAGACAGAGCCGAGAGCCGAGAGTCCGCGACAGACGAGGCGAGAGC 181
Oy 41 GlyGluArgTrpGluLeuAlaLeuGluArgPheTrpAspTrpLeuArgTrpValGlnThr 60
Db 182 GCCAGAGCGCTGGAACTGCGACTGGGTCTTTGGGATTAACCTCGCGCGGTGACAGACA 241
Oy 61 LeuSerGluGluValGluGluGluLeuLeuSerSerGlnValThrGluGluLeuArgAla 80
Db 242 CTGTCTGAGCAGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 301
Oy 81 LeuMetAspGluThrMetLysGluLeuLysValTrpLysSerGluLeuGluGluGluLeu 100
Db 302 CTGATGAGAGAGACCATGAGAGAGTTGAAGGCTTCAAAATCGGAATGAGAGAGACACTG 361
Oy 101 ThrProValAlaGluGluGluThrArgAlaArgLeuSerLysGluLeuGluAlaAlaGlnAla 120
Db 362 ACCCGGAGTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 421
Oy 121 ArgLeuGluAlaAspMetGluAspValLysGlyArgLeuValGlnTrpArgGlyGluVal 140
Db 422 CCGCTGGCGCGAGACATGAGAGAGAGTGTGCGCGCGCTGTGTCACTACCGCGCGAGAGTGTG 481

```

```

RESULT 6
US-08-726-306A-28
Sequence 28, Application US/08726306A
Patent No. 5958684
GENERAL INFORMATION:
APPLICANT: van Leeuwen, Frederik Willem
APPLICANT: Burchach, Johannes Peter Henri
APPLICANT: Grosveld, Franklin G.
TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
NUMBER OF SEQUENCES: 189
CORRESPONDENCE ADDRESS:

```


Oy	141	GlnAlaMwLleuglYInserThgIugluenuarValArqlenuAlasrHslstAry	16
Dd	442	CAGGCATGCTCGGCCAGACACCAGAGACTGGGGTGGCCCTCGCTCCACCTGGC	541
Oy	161	LysLeuArgLySAtgLeuLeuArqAspAlaAspAspleuGlnlyScysLeuAlaValTr	180
Dd	542	AAGCTGCGTAGGGGCTTCCTCCGGATCCGATGACCTGCAGAAAGCCCTGGGAGGTAC	601
Oy	181	GlnAlagLYlaArvgglUgLYlaGluaRgLYleuSerAlalleargluargleu	199
Dd	602	CAGGCCGGGGCCCCGAGGGGCCGAGCGGGCTCAGCCCATCCCGAGCCCTG	658
<hr/>			
RESULT 7			
US-07-849-389-6			
Sequence 6, Application US/07849389			
Patent No. 5525493			
GENERAL INFORMATION:			
APPLICANT: HORNES, Erik			
APPLICANT: UHLEN, Mathias			
TITLE OF INVENTION: CLONING METHOD AND KIT			
NUMBER OF SEQUENCES: 7			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Foley & Lardner			
STREET: 1800 Diagonal Road,			
City: Alexandria			
STATE: Virginia			
COUNTRY: USA			
ZIP: 22313-0299			
COMPUTER READABLE FORM:			
MEDIUM TYPE: floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/07/849,389			
FILING DATE: 19920519			
CLASSIFICATION: 435			
ATTORNEY/AGENT INFORMATION:			
NAME: BENT, Stephen A.			
REGISTRATION NUMBER: 29,768			
REFERENCE/DOCKET NUMBER: 16787/168/DFBC			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: (703)836-9300			
TELEFAX: (703)683-4109			
TELEX: 899149			
INFORMATION FOR SEQ ID NO: 6:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 330 base pairs			
TYPE: NUCLEIC ACID			
STRANDEDNESS: single			
TOPOLOGY: linear			
MOLECULE TYPE: DNA (genomic)			
ORIGINAL SOURCE:			
ORGANISM: HUMAN LIPOPROTEIN E GENE			
FEATURE:			
NAME/KEY: CDS			
LOCATION: 1..330			
US-07-849-389-6			
<hr/>			
Alignment Scores:			
Pred. NO.: 5,47e-41 Length: 330			
Score: 503.00 Matches: 106			
Percent Similarity: 97.27% Conservative: 1			
Best local Similarity: 96.36% Mismatches: 3			
Query Match: 31.66% Indels: 0			
DB: 1 Gaps: 0			
<hr/>			
US-09-827-854-16 (1-317) x US-07-849-389-6 (1-330)			
Oy	87	LysGluLeuLysAlArTrLySserGluLeuGluGluGlnLeuThrProValAlaGluGlu	106
Dd	1	AAGAAGTTGAAGGGCTTAACAATCGAACCTGAGAGAACCACTGACCCCGGTGGGAGAGAG	60

```

US-09-827-854-16 (1-317) x US-08-617-256-24 (1-252)

RESULT 8
US-08-617-256-24
: Sequence 24, Application US/08617256
: Patent No. 6043031
: GENERAL INFORMATION:
: APPLICANT: Kvster, Hubert
: TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
: NUMBER OF SEQUENCES: 33
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHIVE & COCKFIELD
: STREET: 60 State Street, suite 510
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109-1875
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/617,256
: FILING DATE: March 18, 1996
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/406,199
: FILING DATE: March 17, 1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Arnold, Beth A.
: REGISTRATION NUMBER: 35,430
: REFERENCE/DOCKET NUMBER: SQT-013CP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)227-7400
: TELEFAX: (617)227-5941
: INFORMATION FOR SEQ. ID NO: 24:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 252 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-08-617-256-24

Alignment Scores:
Pred. No.: 1,43e-28 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.54% Indels: 1
DB: 3 Gaps: 0

```


QY 109 AlAArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlnAlaAspMetGluAsp 128
Db 2 GCACGGCTGTCCAAGAGCTGCAGAGCGCGCCAGCGCCGCTGGCGCGGCAGCATGTAGAGAC 61
QY 129 ValCysGlyArgLeuValAlaGlnTyrArgGlyGluValAlaGlnAlaMetLeuGlnInserThr 148
Db 62 GTGTGGCC-CGCTGTGTGACAGTACCGCGGAGGTGCAGGCTCAGCTCGGCGCAGACACC 120
QY 149 GAGAGCTGTCCAAGAGCTGCAGAGCGCGCCAGCGCCGCTGGCGCGGCAGCATGTAGAGAC 168
Db 121 GAGAGCTGTCCAAGAGCTGCAGAGCGCGCCAGCGCCGCTGGCGCGGCAGCATGTAGAGAC 180
QY 169 AspaAlaAspAspLeuGlnLysCysLeuAlaValTyrGlnAlaGlnAlaArgGluGlyAla 188
Db 181 GATGCCGATGACCTGCAGAGTCCCTGTGACAGTACAGCGCGCGCGCGCGCGCGCGCC 240
QY 189 GluArgGlyLeu 192
Db 241 GAGCGCGGCTC 252

RESULT 9

US-09-287-141-24
Sequence 24, Application US/09287141
Patent No. 6197498

GENERAL INFORMATION:

APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,141
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA: 08/406,199
APPLICATION NUMBER: 08/406,199

ATTORNEY/AGENT INFORMATION:

NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-5941
TELEFAX: (617)227-7400
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-141-24

Alignment Scores:

Pred. No.: 1.43e-28 Length: 252
Score: 374.00 Matches: 82

Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.54% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-16 (1-317) x US-09-287-141-24 (1-252)

QY 109 AlAArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlnAlaAspMetGluAsp 128
Db 2 GCACGGCTGTCCAAGAGCTGCAGAGCGCGCCAGCGCCGCTGGCGCGGCAGCATGTAGAGAC 61
QY 129 ValCysGlyArgLeuValAlaGlnTyrArgGlyGluValAlaGlnAlaMetLeuGlnInserThr 148
Db 62 GTGTGGCC-CGCTGTGTGACAGTACCGCGGAGGTGCAGGCTCAGCTCGGCGCAGACACC 120
QY 149 GAGAGCTGTCCAAGAGCTGCAGAGCGCGCCAGCGCCGCTGGCGCGGCAGCATGTAGAGAC 168
Db 121 GAGAGCTGTCCAAGAGCTGCAGAGCGCGCCAGCGCCGCTGGCGCGGCAGCATGTAGAGAC 180
QY 169 AspaAlaAspAspLeuGlnLysCysLeuAlaValTyrGlnAlaGlnAlaArgGluGlyAla 188
Db 181 GATGCCGATGACCTGCAGAGTCCCTGTGACAGTACAGCGCGCGCGCGCGCGCGCGCC 240
QY 189 GluArgGlyLeu 192
Db 241 GAGCGCGGCTC 252

RESULT 10

US-09-431-613-24
Sequence 24, Application US/09431613
Patent No. 6221601

GENERAL INFORMATION:

APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/431,613
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA: 08/406,199
APPLICATION NUMBER: 08/406,199

ATTORNEY/AGENT INFORMATION:

NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002G
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-5941
TELEFAX: (617)227-7400
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: CDNA
US-09-431-613-24
Alignment Scores:
Pred. No.: 1,436-28
Score: 374.00
Percent Similarity: 97.62%
Best Local Similarity: 97.62%
Query Match: 23.54%
DB: 4
Length: 252
Matches: 82
Conservative: 0
Mismatch: 2
Indels: 1
Gaps: 0
US-09-827-854-16 (1-317) x US-09-431-613-24 (1-252)
QY 109 ALaArGlEuSerLySGluLeuGlnAlaAlaGlnAlaArGlEuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGCGCGGCGAGCCCGGCTGGCCGACATGGAGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGGC-CGCTGTGTGAGTACCGCGCGAGGTGACAGGCATGCTCGCGACAGACACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGCGGGTGGCTGCTCCACCTGCCCAAGCTCGTAAAGCGCTCTCCGCG 180
QY 169 AspAlaAspAspLeuGlnLysCysLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGTCCCTGGCAGTGTACAGGCGGGGCCCGCGAGGGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTC 252
RESULT 11
US-09-504-245-24
Sequence 24, Application US/09504245
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe LLP
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/504, 245
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617, 256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406, 199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002J
TELECOMMUNICATION INFORMATION:
TELEPHONE: 858-450-8400
TELEFAX: 858-587-5360
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-504-245-24
Alignment Scores:
Pred. No.: 1,436-28
Score: 374.00
Percent Similarity: 97.62%
Best Local Similarity: 97.62%
Query Match: 23.54%
DB: 4
Length: 252
Matches: 82
Conservative: 0
Mismatch: 2
Indels: 1
Gaps: 0
US-09-827-854-16 (1-317) x US-09-504-245-24 (1-252)
QY 109 ALaArGlEuSerLySGluLeuGlnAlaAlaGlnAlaArGlEuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGCGCGGCGAGGTGACAGGCATGCTGGCCAGAGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGGC-CGCTGTGTGAGTACCGCGCGAGGTGACAGGCATGCTCGCGACAGACACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGCGGGTGGCTGCTCCACCTGCCCAAGCTCGTAAAGCGCTCTCCGCG 180
QY 169 AspAlaAspAspLeuGlnLysCysLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGTCCCTGGCAGTGTACAGGCGGGGCCCGCGAGGGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTC 252
RESULT 12
US-09-287-682-24
Sequence 24, Application US/09287682
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287, 682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617, 256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406, 199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002E
TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-682-24

Alignment Scores:
Pred. No.: 1,43e-28 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.54% Indels: 1
DB: Gaps: 0

US-09-827-854-16 (1-317) x US-09-287-682-24 (1-252)

OY 109 AlaArgLeuSerIysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
Db 2 GCACGGCTGTCCAAAGAGCTGCAGCGCGCGCAGCCCGCTGGCGCGGCATGAGGAGGAC 61

OY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
Db 62 GTGTGGCGC-CGCTGTGTGCAAGTACCCGCGGAGGTGACAGGCCATGTCTCGGCCAGACACC 120

OY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
Db 121 GAGGAGCTGGGGGTGGCGCTCCGCTCCACCTGCGCAAGCTGCTAAGCGGCTCTCTCCGC 180

OY 169 AspAlaAspAspLeuGlnIscysLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
Db 181 GATGCGCATGACCTGCAGAAATCCCTGCGCATGTACAGGCGCGGCGCGCGAGGCGCC 240

OY 189 GluArgGlyLeu 192
Db 241 GAGCGGCGCCTC 252

RESULT 13
US-09-287-679-24
Sequence 24, Application US/09287679
Patent No. 6258538
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
APPLICANT: Braun, Andreas
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehman White & Mcauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,679
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-679-24

Alignment Scores:
Pred. No.: 1,43e-28 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.54% Indels: 1
DB: Gaps: 0

US-09-827-854-16 (1-317) x US-09-287-679-24 (1-252)

OY 109 AlaArgLeuSerIysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
Db 2 GCACGGCTGTCCAAAGAGCTGCAGCGCGCGCAGCCCGCTGGCGCGGCATGAGGAGGAC 61

OY 129 ValCysGlyArgLeuValGlnTyrArgGlyGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
Db 62 GTGTGGCGC-CGCTGTGTGCAAGTACCCGCGGAGGTGACAGGCCATGTCTCGGCCAGACACC 120

OY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
Db 121 GAGGAGCTGGGGGTGGCGCTCCGCTCCACCTGCGCAAGCTGCTAAGCGGCTCTCTCCGC 180

OY 169 AspAlaAspAspLeuGlnIscysLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
Db 181 GATGCGCATGACCTGCAGAAATCCCTGCGCATGTACAGGCGCGGCGCGCGAGGCGCC 240

OY 189 GluArgGlyLeu 192
Db 241 GAGCGGCGCCTC 252

RESULT 14
US-09-397-766-24
Sequence 24, Application US/09397766
Patent No. 6268144
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
APPLICANT: Braun, Andreas
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehman White & Mcauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,679
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996

APPLICATION NUMBER: US/09/397,766
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002I
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-397-766-24

Alignment Scores:
Pred. No.: 1,43e-28 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.54% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-16 (1-317) x US-09-397-766-24 (1-252)

QY 109 AlaAgtgLeuSerLygGluLeuGlnAlaGlnAlaAgtgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAAAGAGCTGCAGAGCGGCGCGCCGCTGGCGGACATTGAGAGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGCGC-CGCTGTGTGTCAGTACCGCGCGAGGTGCAGGCCATGCTCGGCCAGAGCAC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGCGGCTCGCTCCACCTGCGCAAGCTCGTAAAGCGGCTCTCCGCG 180

QY 169 AspAlaAspAspLeuGlnLysCysLeuAlaValTyrGlnAlaGlyAlaArgGluGluAla 188
DB 181 GATGCCGATGACCTGCAGAACTCCTGGCAGTGTACCAAGGCGGCGGCCGCGAGGGCGCG 240

QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCGCTC 252

RESULT 15
US-09-287-681-24
Sequence 24, Application US/09287681
Patent No. 6277573
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,681
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-681-24

Alignment Scores:
Pred. No.: 1,43e-28 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.54% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-16 (1-317) x US-09-287-681-24 (1-252)

QY 109 AlaAgtgLeuSerLygGluLeuGlnAlaGlnAlaAgtgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAAAGAGCTGCAGAGCGGCGCGCCGCTGGCGGACATTGAGAGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGCGC-CGCTGTGTGTCAGTACCGCGCGAGGTGCAGGCCATGCTCGGCCAGAGCAC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGCGGCTCGCTCCACCTGCGCAAGCTCGTAAAGCGGCTCTCCGCG 180

QY 169 AspAlaAspAspLeuGlnLysCysLeuAlaValTyrGlnAlaGlyAlaArgGluGluAla 188
DB 181 GATGCCGATGACCTGCAGAACTCCTGGCAGTGTACCAAGGCGGCGGCCGCGAGGGCGCG 240

QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCGCTC 252

Search completed: March 14, 2003, 20:18:40
Job time : 39.7343 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 17:32:58 ; Search time 65.8753 Seconds

(without alignments)
3377.460 Million cell updates/sec

Title: US-09-827-854-16

Perfect score: 1589

Sequence: 1 MKVLMALLVFLAGCAQKV.....VEKVOAVGTSAPVPSDNH 317

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 501302 seqs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp

-O=/cgn2.1/USPTO_spool/US09827854/runat_11032003.101612.27578/app.query.fasta.1.3576

-DB=Published_Applications_NA -OFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1

-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62

-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100

-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0

-MAXLEN=200000000 -USER=US09827854.eccn.1.1.298.runat_11032003.101612.27578

-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MAP -LARGEDUERT -NEG_SCORES=0 -WAIT -LONGLOG

-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published_Applications_NA:

1: /cgn2.6/ptodata/1/pubpna/US07_PUBCOMB.seq:

2: /cgn2.6/ptodata/1/pubpna/PTCT_NEW_PUB.seq:

3: /cgn2.6/ptodata/1/pubpna/US06_NEW_PUB.seq:

4: /cgn2.6/ptodata/1/pubpna/US06_PUBCOMB.seq:

5: /cgn2.6/ptodata/1/pubpna/US07_NEW_PUB.seq:

6: /cgn2.6/ptodata/1/pubpna/US07_PUBCOMB.seq:

7: /cgn2.6/ptodata/1/pubpna/US08_NEW_PUB.seq:

8: /cgn2.6/ptodata/1/pubpna/US08_PUBCOMB.seq:

9: /cgn2.6/ptodata/1/pubpna/US09_NEW_PUB.seq:

10: /cgn2.6/ptodata/1/pubpna/US09_PUBCOMB.seq:

11: /cgn2.6/ptodata/1/pubpna/US10_NEW_PUB.seq:

12: /cgn2.6/ptodata/1/pubpna/US10_PUBCOMB.seq:

13: /cgn2.6/ptodata/1/pubpna/US60_NEW_PUB.seq:

14: /cgn2.6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1589	100.0	1156	10	US-09-827-854-9
2	1582	99.6	1156	10	US-09-827-854-10
3	1577	99.2	1156	9	US-09-870-759-129
4	1577	99.2	1156	9	US-09-802-640-17

5	1577	99.2	1156	10	US-09-827-854-8	Sequence 8, Appl1
6	1577	99.2	1291	12	US-10-044-090-454	Sequence 454, App
7	1573	99.0	1156	10	US-09-827-854-12	Sequence 12, Appl
8	1569	98.7	1156	10	US-09-827-854-11	Sequence 11, Appl
9	1568	98.7	1157	10	US-09-954-456-760	Sequence 760, App
10	1568	98.7	1157	10	US-09-880-107-2244	Sequence 2244, App
11	1565	98.5	1156	10	US-09-827-854-7	Sequence 7, Appl1
12	1368.5	86.1	41907	10	US-09-967-013-5	Sequence 5, Appl1
13	859.5	54.1	786	10	US-09-925-302-133	Sequence 133, App
14	625	39.3	478	10	US-09-964-824A-374	Sequence 374, App
15	625	39.3	478	10	US-09-880-107-2491	Sequence 2491, App
16	458.5	28.9	449	10	US-09-960-352-4726	Sequence 4726, App
17	455	28.6	356	10	US-09-960-352-5420	Sequence 5420, App
18	439.5	27.7	442	10	US-09-960-352-9395	Sequence 9395, App
19	435	27.4	414	10	US-09-960-352-8237	Sequence 8237, App
20	428.5	27.0	416	10	US-09-960-352-8720	Sequence 8720, App
21	422.5	26.6	423	10	US-09-960-352-14047	Sequence 14047, A
22	413	26.0	253	10	US-09-179-536B-130	Sequence 130, App
23	405.5	25.5	409	10	US-09-960-352-9148	Sequence 9148, App
24	401.5	25.3	416	10	US-09-960-352-9797	Sequence 9797, App
25	387.5	24.4	425	10	US-09-960-352-4497	Sequence 4497, App
26	380.5	23.9	390	10	US-09-960-352-1311	Sequence 1311, App
27	379.5	23.9	391	10	US-09-960-352-1278	Sequence 1278, App
28	379.5	23.9	392	10	US-09-960-352-10599	Sequence 10599, A
29	379.5	23.9	401	10	US-09-960-352-1187	Sequence 1187, App
30	375.5	23.6	401	10	US-09-960-352-8042	Sequence 8042, App
31	374	23.5	252	10	US-09-796-416-24	Sequence 24, Appl
32	374	23.5	252	10	US-09-879-341-24	Sequence 24, Appl
33	369.5	23.3	404	10	US-09-960-352-2325	Sequence 2325, App
34	369.5	23.3	405	10	US-09-960-352-14063	Sequence 14063, A
35	351.5	22.1	377	10	US-09-960-352-176	Sequence 176, App
36	350.5	22.1	378	10	US-09-960-352-4914	Sequence 4914, App
37	350.5	22.1	378	10	US-09-960-352-10040	Sequence 10040, A
38	350.5	22.1	386	10	US-09-960-352-10361	Sequence 10361, A
39	348.5	21.9	388	10	US-09-960-352-11966	Sequence 11966, A
40	345.5	21.7	396	10	US-09-960-352-3457	Sequence 3457, App
41	344.5	21.7	377	10	US-09-960-352-10540	Sequence 10540, A
42	338.5	21.3	432	10	US-09-960-352-8934	Sequence 8934, App
43	338	21.3	332	10	US-09-960-352-6552	Sequence 6552, App
44	337.5	21.2	376	10	US-09-960-352-5408	Sequence 5408, App
45	335.5	21.1	373	10	US-09-960-352-3681	Sequence 3681, App

ALIGNMENTS

RESULT 1
US-09-827-854-9
Sequence 9, Application US/09827854
Patent No. US20020123093A1
GENERAL INFORMATION:
APPLICANT: Zannis, Vassilis
TITLE OF INVENTION: Compounds and methods for lowering
TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia
FILE REFERENCE: 07180/004003
CURRENT APPLICATION NUMBER: US/09/827,854
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: US 09/679,088
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 09/544,386
PRIOR FILING DATE: 2000-04-06
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 1156
TYPE: DNA
ORGANISM: Homo sapiens
US-09-827-854-9
Alignment Scores:
Pred. No.: 3.42e-137 Length: 1156
Score: 1589.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-827-854-16 (1-317) x US-09-827-854-9 (1-1156)
QY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 61 ATGAAGCTTCTGGGCTGGTGTGTCACATCTCTGGCAGAGATGCCAGGCAAGTG 120
QY 21 GluGlnAlaValAlaGlnThrGluProGluProGluLeuArgGlnGlnThrGluProGlnSer 40
DB 121 GACCAAGCGGTGGAGACAGACGGCGGAGCCGAGCTCGCCAGACAGCCAGATGGCAGAGC 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlnArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 181 GGGCAGCGGTGGAGACTGGCAGCTGGTGGTGGATTACTGGCGTGGGTCAGAC 240
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 241 CTGTCTGAGCAGGTGCAGAGAGCTGCTCAGCTCCAGGTCCAGCAGCACTGAGGCGG 300
QY 81 LeuMetAspGluThrMetLysGlnLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
DB 301 CTGATGACAGACCACTGAAGAGGTGAAGGCTTACAAATCGGAGCTGGAGACACTG 360
QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 361 ACCCGGTTGGCGGAGAGACGGCGGACGGCTGTCCAGAGAGTGCAGGCGGCGGACGGC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 421 CGGCTGGGCGGACATGTGGAGCTGTGGCGGCGGCTGGTCATACCGCGCGAGGTG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCAATGCTGGCCAGACACCGAGAGAGCTGGGTGGCTCCCTCCCTCCAGCTGGCG 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
DB 541 AACCTGCGTAAAGCGGCTCCCGCGATGCCGATGACCTCGAGAAAGTCCGAGGAGTAC 600
QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 601 CAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
QY 201 ProLeuValAlaGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
DB 661 CCCCTGGTGGACAGGCGGCGGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
QY 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGluMetGly 240
DB 721 CTACAGAGAGCGGCGGCGGCTGGGCGGAGCGGCTGGCGGCGGAGTGGAGATGGGCG 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValAlaArgAlaLys 260
DB 781 AGCGGAGACCGGCGGCGGCTGGAGCGAGGTGAAGAGACAGGTGGCGAGGTGGCGGCGG 840
QY 261 LeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
DB 841 CTGGAGAGAGAGCGGCGGCGGATACGCTGGCAGGCGGAGGCGCTTCCAGGCGGCGGCTC 900
QY 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
DB 901 AGCTGTTGGAGCGGCTGGTGGAGAAATGACAGCGCAGTGGGCGGCGGCTGGTGGAGAG 960
QY 301 ValGlnAlaAlaValGlnThrSerAlaAlaProValProSerAspAsnHis 317
DB 961 GTGCAAGCTGCGGTGGCGACAGCGCGGCGGCTGTGCTCCAGCAGCAATATCAC 1011
RESULT 2
US-09-827-854-10
; Sequence 10, Application US/09827854

; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: zannis, Vassilis
; APPLICANT: Kyriacos, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-10
Alignment Scores:
Pred. No.: 1,5e-136 Length: 1156
Score: 1582.00 Matches: 316
Percent Similarity: 99.68% Conservative: 0
Best Local Similarity: 99.68% Mismatches: 1
Query Match: 99.56% Indels: 0
DB: 10 Gaps: 0
US-09-827-854-16 (1-317) x US-09-827-854-10 (1-1156)
QY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 61 ATGAAGCTTCTGGGCTGGTGTGTCACATCTCTGGCAGAGATGCCAGGCAAGTG 120
QY 21 GluGlnAlaValAlaGlnThrGluProGluProGluLeuArgGlnGlnThrGluProGlnSer 40
DB 121 GACCAAGCGGTGGAGACAGACGGCGGAGCCGAGCTCGCCAGACAGCCAGTGGCAGAGC 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlnArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 181 GGGCAGCGGTGGAGACTGGCAGCTGGGCTTTGGGATTACTGGCGTGGGTCAGAC 240
QY 61 LeuSerGluGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 241 CTGTCTGAGCAGGTGCAGAGAGCTGCTCAGCTCCAGGTCCAGCAGCACTGAGGCGG 300
QY 81 LeuMetAspGluThrMetLysGlnLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
DB 301 CTGATGACAGACCACTGAAGAGGTGAAGGCTTACAAATCGGAGCTGGAGACACTG 360
QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 361 ACCCGGTTGGCGGAGAGACGGCGGACGGCTGTCCAGAGAGTGCAGGCGGCGGACGGC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 421 CGGCTGGGCGGACATGTGGAGCTGTGGCGGCGGCTGGTGGATGATGCCGCGGAGGTG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCAATGCTGGCCAGACACCGAGAGAGCTGGGTGGCTCCCTCCAGCTGGCGG 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
DB 541 AACCTGCGTAAAGCGGCTCCCGCGATGCCGATGACCTCGAGAAAGTCCGAGCTGGAGTAC 600
QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 601 CAGGCGGCGGCGGCGGAGGCGGCGGCGGCGGCTGACGCGCATCCGCGAGCGGCTGGG 660
QY 201 ProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220


```
Db 661 CCCCTGCTGGACAGGCGCGCTGCGGGCCGCGACTGTGGCTCCCTGGCGCGCCAGCCG 720
QY 221 LeuGlnGluArgAlaGlnAlaIleuValThrPheLeuAlaGlyCysGlnAlaIleuVal 240
Db 721 CTACAGAGAGCGGCGCCAGGCGCTGGCGAGCGGCTGCGCGCGGATGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValIleuGlnGlnAlaGlnAlaIleuVal 260
Db 781 AGCGGAGACCGCGACCGCTGGAGCGAGGTGAAGAGACAGTGGCGGAGGTGGCGCGCAG 840
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuIys 280
Db 841 CTGGAGAGACAGCGCCAGAGATAGCCCTGACAGCGCGAGCGCTCCAGCGCGCTCAG 900
QY 281 SerTyrPheGluProLeuValGluAspMetGlnArgGlnIleuValGlnIleuValGlnIys 300
Db 901 AGCTGTTGAGACCCCTGGTGTGAGACATGCAAGCGCCAGTGGCGCGGCTGGTGGAGAG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGAGCTGCCGTGGGACACAGCGCGCCCTGTGTCACAGGACATAC 1011

RESULT 3
US-09-870-759-129
; Sequence 129, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; PRIORITY FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIORITY FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 129
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(1014)
; OTHER INFORMATION:
US-09-870-759-129

Alignment Scores:
Pred. No.: 4,31e-136
Score: 1577.00
Percent Similarity: 99.68%
Best Local Similarity: 99.68%
Query Match: 99.24%
DB: 9 Gaps: 0

US-09-827-854-16 (1-317) x US-09-870-759-129 (1-1156)
QY 1 MetLysValLeuThrAlaIleuValIleuValThrPheLeuAlaGlyCysGlnAlaIleuVal 20
Db 61 ATGAAAGTTCGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGlnIleuVal 40
Db 121 GAGCAAGCGGTGGAGACAGCGGAGCGCGGAGCTGGCGGACAGACAGCGAGTGGCGAGAC 180
QY 41 GlyGlnArgTyrGluLeuAlaLeuGlnIleuValThrPheTyrPheTyrLeuArgTyrValGlnThr 60
Db 181 GGCCAGCGCTGGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 61 LeuSerGlnGlnValGlnGlnIleuLeuSerGlnValThrGlnGlnIleuValAla 80
Db 241 CTGTCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
QY 81 LeuMetAspGluThrMetLysGluLeuIleuValIleuValIleuValIleuValIleuVal 100
```

```
Db 301 CTGATGAGACGAGACCATGAGAGAGTTGAAGCGCTCAATTCGAACTGAGAGAACAACTG 360
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGlnIleuGlnAlaIleuVal 120
Db 361 ACCCGGTGGCGAGAGAGCGGGGACGCTGTCCMAAGACTGAGGCGGCGCGAGGCG 420
QY 121 ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnIleuValIleuVal 140
Db 421 CGGCTGGCGCGGAGACATGAGAGACGTGTGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGT 480
QY 141 GlnAlaMetLeuGlyIleuSerThrGlnIleuValArgValAlaArgLeuAlaSerHisLeuArg 160
Db 481 CAGCCCATCTCGCGCCAGAGCAGCGAGAGCTGCGGGTGTGCGCTGCGCTGCCACCTGCG 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 541 AAGCTGCTTAACCGGCTCTCCGCGATGCGCATGACCTGCAGAAAGCGCTGCGCATGTAC 600
QY 181 GlnAlaGlyAlaArgGluGlnIleuValArgGlyLeuSerAlaIleuArgLeuGlnIleuVal 200
Db 601 CAGCGCGGCGCGCGGAGGCGCGGCGCGGCGGCGCTTCAGCGCATCCGAGCGCGCTGGGG 660
QY 201 ProLeuValGluGlnGlyArgValAlaArgAlaIleuValIleuValIleuValIleuVal 220
Db 661 CCCCTGTGGACAGAGGCGCGTGGCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 221 LeuGlnGluArgAlaGlnAlaIleuValIleuValIleuValIleuValIleuValIleuVal 240
Db 721 CTACAGAGAGCGGCGCGGAGCGCTGGGCGGAGCGCGCGCGGCTGCGGATGAGTGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValIleuGlnGlnValAlaGlnValAlaArgAlaIys 260
Db 781 ACCCGGAGCGCGAGCGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
QY 261 LeuGlnGluGlnAlaGlnGlnIleuArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuIys 280
Db 841 CTGGAGAGACAGCGCCAGAGATAGCCCTGACAGCGCGAGCGCTTCAGAGCGCGCGCTCAG 900
QY 281 SerTyrPheGluProLeuValGluAspMetGlnArgGlnIleuValGlnIleuValGlnIys 300
Db 901 AGCTGTTGAGACCCCTGGTGTGAGACATGACATGACAGCGCCAGTGGCGCTGTGAGAG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGAGCTGCCGTGGGACACAGCGCGCCCTGTGTCACAGGACATAC 1011

RESULT 4
US-09-802-640-17
; Sequence 17, Application US/09802640
; Publication No. US20030036057A1
; GENERAL INFORMATION:
; APPLICANT: Braun, Andreas
; APPLICANT: Bonsal Aruna
; APPLICANT: Kleyen Patrick
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND THEIR USE
; FILE REFERENCE: 24736-2048
; CURRENT APPLICATION NUMBER: US/09/802,640
; PRIORITY FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(1014)
; OTHER INFORMATION: Nucleotide Sequence encoding apolipoprotein E
; OTHER INFORMATION: (APOE)
US-09-802-640-17
```


Alignment Scores:		
Pred. No.:	4,316-136	1156
Score:	1577.00	316
Percent Similarity:	99.68%	Conservative: 0
Best Local Similarity:	99.68%	Mismatches: 1
Query Match:	99.24%	Indels: 0
BB:	9	Gaps: 0

US-09-827-854-16 (1-317) x US-09-802-640-17 (1-1156)

QY	1	MeuysValLeuTPRLAAlaIleuLeuValThrhPheLeuNlaIcyssGlnAlaIysVal	20
Db	61	ATGAGGCTCTGTGGCTGGCTGGTGTACATCTCGGACGATGCCAGGCCAAGGTG	120
QY	21	GluGlnAlaValGluThrGluProGluProGluIleuArgGlnGlnThrGluTrpGlnSer	40
Db	121	GAGCAACCGGTGGAGACAGAACCCGAGCCGAGCTGGCCACGACACGACTGGCAGAC	180
QY	41	GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr	60
Db	181	GGCCAGCGCTGGGAACAGGCACTGGGTGGCTTTGGGATTAACCTGGCTGGTCGACAGA	240
QY	61	LeuSerGluGlnValGlnGlnGluIleuLeuSerSerGlnValThrhGlnGluLeuArgAla	80
Db	241	CTGTCTGACAGGTGCAGAGAGAGACTGTCAAGCTCCAGGTCAACCAAGAACTAGAGGGC	300
QY	81	LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGluGlnLeu	100
Db	301	CTGATGGACGAGACCATGATGAAGAGTTAAGGCTTACAAATCGGAACTGGAGAAACAAC	360
QY	101	ThrProValAlaGluGluThrThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	361	ACCCGGTGGGGAGAGACGCGGGCCAGGCTGTCCAAAGAGCTGCAGCGCGCGAGGCC	420
QY	121	ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal	140
Db	421	CGCTGGGGCGGGACATGAGAGACTGTGGCGCGCTGTGTGAGATGCCGCGGAGGTG	480
QY	141	GlnAlaMetLeuGlnGlyGlnSerThrhGlnGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	481	CAGGCCATGCTCGGGCCAGAGACCGAGAGACTGGGGTGGCTCGCTCCACACTGGCG	540
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTrp	180
Db	541	AAGCTGGCTAAGCGCGCTCTCCGCGAATGCCGATGACTCGCAAGAGCGCTCGCAGCTGAC	600
QY	181	GlnAlaGlyAlaArgGluGluGluAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200
Db	601	CAGGCCGGGGCGCCGAGGGGCGCGAGCGGGCTCAAGCCCATCCGCGAGCGCTGGGG	660
QY	201	ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro	220
Db	661	CCCTGGTGGAAACAGGGCCGGTGGGGCGCCACTGTGGCTCCGTGGCGGGCCAGCGG	720
QY	221	LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly	240
Db	721	CTACAGAGCGGGGCCAGGCTGTGGGGAGACGGCTGGCGCGCGGATGTGAGAGATGGGC	780
QY	241	SerArgThrArgAspArgLeuAspGluValLysGluGlnAlaAlaGluValArgAlaLys	260
Db	781	AGCCGAGCCCGCACCCCTGTGAGAGAGTAAAGAGAGAGGTGGCGGAGGTGCGCGCAAG	840
QY	261	LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluLysPheGlnAlaArgLeuLys	280
Db	841	CTGGAGGAGCAGGCCACAGCATACGCTTCAGGCCGAGGCTTCCAGGCCCGCCCTCAAG	900
QY	281	SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys	300
Db	901	AGCTGTTGACGCCCTCGTGTGAAGACATCCACGCCCAGTGGGGCGGCTGTGTGAGAAAG	960
QY	301	ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis	317
Db	961	GTCAGAGCTCCGTCGGGACACAGCGGCCCTGTGTGCCAGCACAATAC	1011

```

RESULT 5
US-09-827-854-8
: Sequence 8, Application US/09827854
: Patent No. US20020123093A1
: GENERAL INFORMATION:
: APPLICANT: Zannis, Vassilis
: TITLE OF INVENTION: Compounds and methods for lowering
: TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia
: FILE REFERENCE: 07180/004003
: CURRENT APPLICATION NUMBER: US/09/827,854
: CURRENT FILING DATE: 2001-04-05
: PRIOR APPLICATION NUMBER: US 09/679,088
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 09/544,386
: PRIOR FILING DATE: 2000-04-06
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 8
: LENGTH: 1156
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-827-854-8

```

Alignment Scores:	
Pred. No.:	4,31e-136
Score:	1577.00
Percent Similarity:	99.688
Best Local Similarity:	99.688
Query Match:	99.248
DB:	10
US-09-827-854-16 (1.317) x US-09-827-854-8 (1.1156)	1156
	Matches: 316
	Conservative: 0
	Mismatches: 1
	Indels: 0
	Gaps: 0

Qy	1	MelcysValleuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	61	ATGAGGCTCTGTGGGCTGTGGTCTGGTGCACATTCTCTGGCAGAGATGCGACGCCAAAGTG	120
Qy	21	GlnGlnAlaValGluThrGluTrpGluTrpGluLeuArgGlnGlnThrGluTrpGlnSer	40
Db	121	GAGCAACGGGTGGAGACAGAGCCGGAGCCGAGCTGGCCACAGCAGCAGATGGGCAAGC	180
Qy	41	GlyGlnArgTrpPheLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr	60
Db	181	GGCCAGCGCTGGGAAGTGGCAGCTGCTTTTGGATTACCTGCGCTGGGTGCACACA	240
Qy	61	LeuSerGlnGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80
Db	241	CTGTCTGAGCAGGTGCAGGAGGAGGCTGTCTACACTCCAGAGTCACCAAGACTAGGGCG	300
Qy	81	LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu	100
Db	301	CTGATGACAGACAGACCAATGAAAGATTGAAAGCCCTCAATGCATCGAAGTGAAGAACCACTG	360
Qy	101	ThrProValAlaGluGluTrpThrArgAlaAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	361	ACCCGGGTGGCGGAGAGAGAGCGGGGACGGCTGTCCAAGAGAGCTGCACGGCGGCCAGGGCC	420
Qy	121	ArgLeuGlnAlaAspMetGluAspValLysGlyArgLeuValGlnTyrArgGlyGlnVal	140
Db	421	CGGCTGGGCGCGGACATGAGAGACGTGTGGGCGCCCTGTGTCAATACCGCGGGGAGAGTG	480
Qy	141	GlnAlaMetLeuGlnLysSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	481	CAGGCCACTGCTCGGCCAGAGACACAGAGAGAGCTGGGGGTGGCCCTCCACTCCACTGGCG	540
Qy	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr	180
Db	541	AAGTCGCTAAGGGGCTCTCCGGCAGTCCGAGTACCTGCAGAAAGGGCTGGCAGGTATAC	600
Qy	181	GlnAlaGlyAlaArgGluGlnAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200

QY 181 GlnAlaGlyAlaArgGluGlnGlyAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
Db 601 CAGGGCGGGGGCCGGAGGGGGCCGAGCGGGCCCTCAGCGCATCTCCGGAGGGCCCTGGGG 660
QY 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
Db 661 CCCCTGTGTGAACAGAGGGCGCGCTGGGGCCGCGCACTGTGGGTCCCTGGCGCGGCAAGCG 720
QY 221 LeuGlnGlnArgAlaGlnAlaThrPylGlnArgLeuArgAlaArgMetGlnGluMetGly 240
Db 721 CTACAGGAGGGGGCCAGGCGCTGGGGCGAGCGGCTGGCGCGGATGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
Db 781 AGCGGAGACCGCGAGCGCGCTGGAGAGGTGAAGAGCAGGTGGCGAGAGTGGCCCAAG 840
QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 841 CTGGAGGAGAGCGCCAGCAGATACGCTCGAGCGCGAGCGCTTCAGGCGCGCTCAAG 900
QY 281 SerTrpPheGlnProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
Db 901 AGCTGCTCGAGCGCCCTGTGTGAACATGACAGCGCCAGTGGGCGGGCTGGTGAAGAG 960
QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGGCGTGGGCACACAGCGCCCGCTGTGGCCAGAGCAATACAC 1011

RESULT 9
US-09-954-456-760; Sequence 760, Application US/09954456
; Patent No. US20020115057A1

GENERAL INFORMATION:

; APPLICANT: Young, Paul

; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand

; FILE REFERENCE: 689290-76

; CURRENT APPLICATION NUMBER: US/09/954,456

; PRIOR FILING DATE: 2001-09-18

; PRIOR APPLICATION NUMBER: US/60/233,617

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: US/60/234,052

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: US/60/234,923

; PRIOR FILING DATE: 2000-09-25

; PRIOR APPLICATION NUMBER: US/60/235,134

; PRIOR FILING DATE: 2000-09-25

; PRIOR APPLICATION NUMBER: US/60/235,637

; PRIOR FILING DATE: 2000-09-26

; PRIOR APPLICATION NUMBER: US/60/235,638

; PRIOR FILING DATE: 2000-09-26

; PRIOR APPLICATION NUMBER: US/60/235,711

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,720

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,840

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,863

; PRIOR FILING DATE: 2000-09-27

; NUMBER OF SEQ ID NOS: 2276

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 760

; LENGTH: 1157

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-954-456-760

Alignment Scores:

Pred. No.: 2,88e-135

Score: 1568.00

Percent Similarity: 99.05%

Best Local Similarity: 99.05%

Query Match: 98.68%

Length: 1157

Matches: 314

Conservative: 0

Mismatch: 3

Indels: 0

DB: 10 Gaps: 0
US-09-827-854-16 (1-317) x US-09-954-456-760 (1-1157)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAGCTTCTGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTG 121
QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnProGlnProGlnProGlnProGlnProGln 40
Db 122 GAGCAAGGGGGTGGAGACAGAGCCGAGCGCCAGCTGCGCCAGCAGCAGCAGCAGCAGCAGCAG 181
QY 41 GlyLysArgTrpGlnLeuAlaLeuGlnValArgPheTrpAspTrpLeuArgTrpValGlnThr 60
Db 182 GGGCAGCGCTGGGAACCTGGCACTGGGTGGCTTTTGGGATTACCTGCGCTGGGTGGAGACA 241
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db 242 CTGTCTGAGCAGGTGACAGAGAGAGCTGCTCAGCTCCCAAGTCCACCAAGAACTGAGGCGC 301
QY 81 LeuMetAspGlnTrpMetLysGlnLeuLysAlaTrpLysSerGlnLeuGlnGlnLeu 100
Db 302 CTGATGACAGAGACCATGATGAGAGATTGAAGGCTTCAATGAGGAGCTGAGAGAACTGTG 361
QY 101 ThrProValAlaGlnGlnTrpArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
Db 362 ACCCGGTAGCGAGAGACACCGCGGACGGCTGTCCAAAGAGCTGACAGCGCGAGGCC 421
QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlnArgLeuValGlnTrpArgGlyGlnVal 140
Db 422 CGGCTGGCGCGGACATGAGAGAGAGCTGTGGCGCCCTGTGTGACATGACCGCGGAGAGTG 481
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValAlaGlnAlaSerHisLeuArg 160
Db 482 CAGGCAATGCTGGCGCAGAGACAGGAGAGCTGGCGGTGGCTGCTCCACCTGTGGC 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValLys 180
Db 542 AAGCTGCTTAAGCGGCTCTCCGCGATCCGATGACTGACAGAGCGGCTGGCACTGTAC 601
QY 181 GlnAlaGlnAlaArgGlnGlnAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
Db 602 CAGGCGGGGGCGCGGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 661
QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
Db 662 CCCCTGTGTGAACAGAGGGCGGTGGGGCGCCACTGTGGCTCTGGCGGCGGCGGCGGCGGCGG 721
QY 221 LeuGlnGlnArgAlaGlnAlaThrPylGlnArgLeuArgAlaArgMetGlnGluMetGly 240
Db 722 CTACAGGAGGGGGCCAGGCGCTGGGGCGAGGGCGGTGGCGGGGATGGAGAGATGGGC 781
QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
Db 782 AGTCGAGACCGCGAGCGCGCTGGAGAGGTGAAGAGCGGTGGCGGAGAGTGGCGGCCCAAG 841
QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 842 CTGGAGGAGAGCGCCAGCAGATACGCTGCAAGGCGGAGGCTTCAGAGCGGCGGCTCAAG 901
QY 281 SerTrpPheGlnProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
Db 902 AGCTGCTCGAGCGCCCTGTGTGAACATGACAGCGCCAGTGGGCGGGCTGGTGAAGAGAG 961
QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGCAGGCTGGCGTGGGCACACAGCGCCCGCTGTGGCCAGAGCAATACAC 1012

RESULT 10

US-09-880-107-2244

; Sequence 2244, Application US/09880107

; Patent No. US20020142981A1

; GENERAL INFORMATION:

Db 336 ATGAAGTTCCTGTCGGCGCTGCTGTCACATTCTGCGAGAGATGCCAGGCCAAGGTG 337
QY 21 GtuglnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
|||||
Db 336 GAGCAAGCGGTGGAGACAGACGCCGAGCCNGACTGCGCCAGCAGACCACTGGCAGAGC 277
QY 41 GtuglnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
276 GGGCAGCGCTGGAACTGGCACTGGTGTCTTTGGATTTACCTGCGCTGGGTGAGACA 217
Db 61 LeuSerGluGlnValGlnGluLeuLeu-SerSerGlnValThrGlnGluLeuArgAl 80
216 CTGCTGTGAGCAGGTGCAGAGAGAGCTGCTGAGCTCCAGGTCCACCCAGAACTGAGGGC 157
QY 80 aLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLe 100
Db 156 GCTGATGACAGACCACTGAAGAGAGTTGAAGCCCTACAAATCGGAACGTGGAGAACT 97
QY 100 uThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAl 120
Db 96 GACCCCGGTGGCGGAGAGAGCGGGCGAGCTGTCCAGAGAGCTGCAGCGCGCGAGN 37
QY 120 aArgLeuGlyAlaAspMetGluAspVal 129
Db 36 CCGGCTGGCGCGGACATGAGAGAGCTG 9

RESULT 15

US-09-880-107-2491/C
: Sequence 2491, Application US/09880107
: Patent No. US20020142981A1
: GENERAL INFORMATION:
: APPLICANT: Horne, Darci T.
: APPLICANT: Vockley, Joseph G.
: APPLICANT: Scherf, Uwe
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
: FILE REFERENCE: 44921-5028-NO
: CURRENT APPLICATION NUMBER: US/09/880,107
: PRIORITY FILING DATE: 2001-06-14
: PRIORITY APPLICATION NUMBER: US 60/211,379
: PRIORITY FILING DATE: 2000-06-14
: PRIORITY APPLICATION NUMBER: US 60/237,054
: PRIORITY FILING DATE: 2000-10-02
: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2491
: LENGTH: 478
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Genbank Accession NO. US20020142981A1 N33009
: LOCATION: (1)..(478)
: OTHER INFORMATION: n = a or c or g or t
US-09-880-107-2491

Alignment Scores:

Pred. No.: 2,72e-49 Length: 478
Score: 625.00 Matches: 126
Percent Similarity: 97.69% Conservative: 1
Best Local Similarity: 96.92% Mismatches: 2
Query Match: 39.33% Indels: 1
DB: 10 Gaps: 0

US-09-827-854-16 (1-317) x US-09-880-107-2491 (1-478)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
Db 336 ATGAAGTTCCTGTCGGCGCTGCTGTCACATTCTGCGAGAGATGCCAGGCCAAGGTG 337
QY 21 GtuglnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
|||||
Db 336 GAGCAAGCGGTGGAGACAGACGCCGAGCCNGACTGCGCCAGCAGACCACTGGCAGAGC 277

QY 41 GtuglnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 276 GGGCAGCGCTGGAACTGGCACTGGTGTCTTTGGATTTACCTGCGCTGGGTGAGACA 217
QY 61 LeuSerGluGlnValGlnGluLeuLeu-SerSerGlnValThrGlnGluLeuArgAl 80
Db 216 CTGCTGTGAGCAGGTGCAGAGAGAGCTGCTGAGCTCCAGGTCCACCCAGAACTGAGGGC 157
QY 80 aLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLe 100
Db 156 GCTGATGACAGACCACTGAAGAGAGTTGAAGCCCTACAAATCGGAACGTGGAGAACT 97
QY 100 uThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAl 120
Db 96 GACCCCGGTGGCGGAGAGAGCGGGCGAGCTGTCCAGAGAGCTGCAGCGCGCGAGN 37
QY 120 aArgLeuGlyAlaAspMetGluAspVal 129
Db 36 CCGGCTGGCGCGGACATGAGAGAGCTG 9

Search completed: March 15, 2003, 03:04:20
Job time : 79.8753 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:09:47 ; Search time 1289.69 seconds

(without alignments)
3980.771 Million cell updates/sec

Title: US-09-827-854-16

Perfect score: 1889

Sequence: 1 MKVLMAALVFLAGCOAKV.....VEKVOAVGTSAPVPSDNH 317

Scoring table:

BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_p2n.model -DEV=xlp
-O=/cgn2.1/USPFO/spool/US09827854/runat_11032003_101610_27495/app_query.fasta.1.3576
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=BIOSUM62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=us09827854_gcg.1.1_8826_grunat_11032003_101610_27495 -NCPU=6 -ICPU=3
-NO.XLPRX -NO.MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlm:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1282.5	80.7	1027	13	B1670367
2	1246	78.4	842	12	B6763371
3	1240	78.0	933	13	B1668318
4	1212	76.3	938	12	B6761746
5	1209.5	76.1	942	13	B1600906
6	1195	75.2	800	13	BM042094
7	1194.5	75.2	927	12	BG472299
8	1191.5	75.0	922	13	B1597743
9	1182.5	74.4	817	12	BG774871
10	1181	74.3	811	13	B1600563
11	1167.5	73.5	845	12	BG829472
12	1144	72.0	790	12	BG707147
13	1141	71.8	919	13	B1551475
14	1131	71.2	757	13	BM042228
15	1121.5	70.6	907	12	BG706129
16	1120	70.5	706	14	BM728696
17	1120	70.5	741	12	BG762924
18	1118	70.4	1110	11	AK010261
19	1116	70.2	804	12	BG702752
20	1115	70.2	834	13	BM042676
21	1112.5	70.0	808	13	B1668329
22	1109	69.8	855	13	B1616362
23	1106	69.6	803	13	B1670350
24	1103	69.4	782	12	BG716776
25	1103	69.4	812	13	B1601551
26	1103	69.4	965	14	B0677266
27	1100	69.2	757	13	B1603658
28	1097.5	69.1	797	12	BG715366
29	1097	69.0	846	13	B1159757
30	1096	69.0	756	13	BM042153
31	1091	68.7	757	12	BG707750
32	1091	68.7	794	12	B1601279
33	1091	68.7	954	13	BE967543
34	1089	68.5	790	13	B1551066
35	1088	68.5	790	13	B1551811
36	1084.5	68.3	798	12	BG708414
37	1083.5	68.2	891	13	B1549292
38	1077.5	67.8	802	13	B1458355
39	1076	67.7	1100	14	BM914382
40	1075.5	67.7	914	13	B1603523
41	1075	67.7	748	13	B1553085
42	1074.5	67.6	930	12	BF967857
43	1070.5	67.4	812	12	BG769968
44	1070	67.3	888	13	B1544886
45	1069.5	67.3	803	12	BG709360

ALIGNMENTS

RESULT 1
B1670367
LOCUS 603292738P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5312024 5',
DEFINITION B1670367
ACCESSION B1670367
VERSION B1670367.1 GI:15584600
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1027)
NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE

JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1790 row: 1 column: 09
High quality sequence stop: 843.
Location/Qualifiers
1. 1027
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 194 a 308 c 400 g 125 t
ORIGIN
Alignment Scores:
Pred. No.: 6.3e-125 Length: 1027
Score: 1282.50 Matches: 291
Percent Similarity: 90.558 Conservative: 6
Best Local Similarity: 88.728 Mismatches: 20
Query Match: 80.714 Indels: 12
DB: 13 Gaps: 2
US-09-827-854-16 (1-317) x B1670367 (1-1027)
QY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
DB 26 ATGAAGTTCTGTGGCTGTGTGTGCATCTCTGCAGAGATGCAGGCCAAGT 85
QY 21 GluGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
|||||
DB 86 GACCAAGCGGTGAGACAGAGCCGAGCCGACCTGCCAGACAGCCAGTGGCAGAGC 145
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
DB 146 GGCACAGCGGTGAGACAGAGCCGAGCCGACCTGCCAGAGATGCAGGCCAAGT 205
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
DB 206 CTGTCTGACAGAGTGCAGAGGAGTGTCTACCTCCAGGTCCACGAGAACTGAGAGGGG 265
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluGlnGluGlnLeu 100
|||||
DB 266 CTGATGAGACAGACCATGAAAGAGTTGAAGCCTTACAAATCGAATGGAGGACAACTG 325
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluGlnAlaAlaGlnAla 120
|||||
DB 326 ACCCGCGTGGCGAGAGAGACGCGGCGAGCTGTCAAGAGAGTGAAGCGGCGAGGCC 385
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
DB 386 CGGCTGGGCGGACATGAGAGACGTGTGCGGCCCTGTGACATACCGCGGCGAGGTG 445
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||

DB 446 CAGGCCATGCTCGGCCAGAGACCCGAGAGCTCGGGGTGCGCTCCGCTCCACCTGCCG 505
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
|||||
DB 506 AAGCTGCGTAAGGGGCTCTCCGCGATGCCGATGACCTGAGAAAGCGCTGGCAGGTAC 565
QY 181 GlnAlaGlyAlaArgGluGluGluArgGlyLeuSerAlaLeuArgGluLeuGln 200
|||||
DB 566 CAGCGCGGGGCGCCGAGAGCGCGGAGCGCGCTCAGCGCCATCCGACGCGCTGGGG 625
QY 200 YProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGln 220
|||||
DB 626 CCCCCTGTGTGAACAGAGCTCGCGTGGCGCGCCGCTGAGGCTCTGCGCGCGCGCC 685
QY 220 OleuGlnGluArgAlaGlnAlaTrpGlyGluArg-LeuArgAlaArgMetGluGlu-Met 239
|||||
DB 686 GCTACAGAGAGCGGCGCAAGCGCTGGGCGAGCGCTTGGCGCGCGAGTGAAGAGAG 745
QY 240 GlySerArg-ThrArgAspArgLeuAspGluValLysGluGlnValAlaGlu-ValArg 259
|||||
DB 746 GGCAGCGCGGAGCCCGACCGCTGAGCCAGGTGAAGAGACAGTGTGGCGAAGTTCGCG 805
QY 259 LAlaLysLeu---GluGlnGlnAlaGlnGln-ThrArgLeuGln-AlaGluAlaPheGlnAl 277
|||||
DB 806 CCCAAGCTGAGCGAGCCAGGCCAGCCAGCAAGATACGCTCAGAGGCCGAGCTTCTAGGC 865
QY 277 ArgLeuLysSerTrpPheGluPro-----LeuValGluAspMetGlnArgGlnTrpAl 295
|||||
DB 866 CCGGCTCAGAGAGCTGTGGAACCCCTCGGTGGAACGACATGGCGCGCAGTGGGCG 925
QY 295 ArgLeuValGluLysVal-GlnAlaAlaValGlyThrSerAlaAlaProValProSerA 315
|||||
DB 926 CGGCTGTGTGAAMAAAGTCCAGAGCTGCGGGGCGACACCCGCGCTGTGTCCAGC-G 984
QY 315 sPAsnHis 317
|||||
DB 985 ACNATCAT 992
RESULT 2
LOCUS BG763371 842 bp mRNA linear EST 15-MAY-2001
DEFINITION 602735433r1 NIH_MGC_49 Homo sapiens CDNA clone IMAGE:4860585 5',
ACCESSION BG763371
VERSION BG763371.1 GI:14074024
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 842)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DC/DT/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1718 row: b column: 10
High quality sequence stop: 817.
Location/Qualifiers
1. 842
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"

Db	195	GGCCAGCGCTGGGAAC	TGCACTGGATGGGTCCTTTGGGATTACCTGGCGCTGGGTGCAGACA	254
QY	61	LeuSerGIUGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80	
Db	255	CGTGTGAGACAGTGTACAGAGAGAGCTGTCTACACTCCACAGTCCACCAGAACCTGAGGGCG	314	
QY	81	LeuMetAspGluThrMetLeuSerGluLeuLeuValATATATLeuSerGluLeuGlnGlnLeu	100	
Db	315	CTGATGTACAGACACATGTAGAGAGTTGAAGGCGCTTACAAATGTGGAACTGGAGGAACAACATG	374	
QY	101	ThrProValAlaGlnGlnGlnThrArgAlaArgLeuSerLeuSerGluLeuGlnAlaAlaGlnAla	120	
Db	375	ACCCGGTGGCGGAGAGAGACGGGGCAGCGGTCTGTCCAAAGACACTCAGCGGGCGCAGGGCG	434	
QY	121	ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal	140	
Db	435	CGGCTGGGGCGCGACATGAGAGAGCTGTGCGCCCGCTGTGTACAGTACCGCGCGAGGTG	494	
QY	141	GlnAlaMetLeuGlnGlnSerThrGlnGlnGluLeuArgValArgLeuAlaSerHisLeuArg	160	
Db	495	CAGGCCATGCTCTGGCCAGAGACACCGAGAGCTGCGGGTGTGCTGCTCCCACTCGCC	554	
QY	161	LysLeuArgItyArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr	180	
Db	555	AAGTGTGCGTAAAGCGGCTCTCCGCGATGGCCATGTACACTGACAGAAGCGCTGGCAGATGAC	614	
QY	181	GlnAlaGlyAlaArgGlnGlnGlnAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly	200	
Db	615	CAGGCGGGGGCGCGGAGGGCGCCAGAGCGCGCTCTACGCGCCATCCGCGAGCGCTGGGG	674	
QY	201	ProLeuValGlnGlnGlnValArgValArgAlaIleThrValGlySerLeuAlaGlyGlnPro	220	
Db	675	CCCTGTGTGGAACAGAGCGCGCTGGCGGGCGCCACTGTGGTCTCCCTGGCGGCAAGCGG	734	
QY	221	LeuGlnGlnArgAlaGlnAlaIleArgGlyGlnArgLeuArgAlaArgMetGlu---GluMet	239	
Db	735	CTACAGAGAGCGGGGCCAGAGCCCTGGGGCGAGCGGCTGTGCGCCGCGATTTGAGGACACATTG	794	
QY	239	TGlySerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArg	259	
Db	795	GGGACGCGGAGCCCGCAACGCTGTGAACGAGTGAAGGACGAGGTCGGAAGTTTGGCCG	854	
QY	259	IaLysLeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgL	279	
Db	855	GCCAAAGGCTGGGGGACACAGGCGCCGCAATTCGTTGTACAGGCCAA-GCCTTCCAGGCGCGCT	913	
QY	279	eulYsserTrrPheGlu 284		
Db	914	CCAAAGAGCTGTTCAA 930		
RESULT 4				
LOCUS	BG761746	938 bp	mRNA linear EST 15-MAY-2001	
DEFINITION	602711942P1 NIH_MGC_49 Homo sapiens CDNA clone IMAGE:4841411 5',			
ACCESSION	BG761746		mRNA sequence.	
VERSION	BG761746.1	GI:14072399		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
TITLE	NIH-MGC http://mgc.ncl.nih.gov/.			
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)			
COMMENT	Unpublished (1999)			
	Contact: Robert Strausberg, Ph.D.			
	Email: cga@ds-r@mail.nih.gov			
	Tissue Procurement: ATCC/CDT/DRP			
	CDNA Library Preparation: Ling Hong/Rubin Laboratory			
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)			
	Clone sequencing by: Incyte Genomics, Inc.			
	Clone distribution: MGC clone distribution information can be			

found through the I.M.A.G.E.E. Consortium/LEUL at: http://image.lm1.gov Plate: LUCM1674 row: C column: 12 High quality sequence stop: 767. Location/Qualifiers	
1. .938 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_id="NIH_MGC_49" /tissue_type="melanotic melanoma, high MDR (cell line)" /lab_host="DH10B (phage-resistant)" /note="Organ: skin; Vector: pOTB7; site 1: XhoI; site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCCAGAG(g). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."	
BASE COUNT	171 a 272 c 374 g 121 t
ORIGIN	
Alignment Scores:	
Score: No.:	1.52e-117 938
Percent Similarity:	1212.00 Matches: 271
Best Local Similarity:	92.28% Conservative: 4
Query Match:	90.94% Mismatches: 12
	76.27% Indels: 11
DB:	12 Gaps: 2
US-09-827-854-16 (1-317) x BG761746 (1-938)	
Oy	1 MettysValleuTrrPALaAlaLeuLeuValThrrPheLeuAlaGlyCysGlnAlaLysVal 20
Db	57 ATGAAGGTTCTGTGGCGCTGCTGTGTGCATTTCCGCGAGGATGCCAGGCCAGGTG 116
Oy	21 GluGlnAlaValAGluThrGluProGluProGluLeuArgGlnGlnThrrGluTrpGlnSer 40
Db	117 GAGCAAGCGGTGGAGACAGACC CGGAGCCGCAAGCTGCGCCAGCAGACCGAAGTGGCAGAGC 176
Oy	41 GlyGlnArgTrpGluLeuAlaLeuGluLysrgherrPasprryTrleuArgTrpValGlnThr 60
Db	177 GCCCGCCCTGGGAACCTGGCACTGGGTGGCTTTGGGATrrACCTGGCTGGGTCCAGCA 236
Oy	61 LeuserGluGlnValAGlnGluLeuLeuserSerGlnValThrrGlnGluLeuArgLa 80
Db	237 CTGTCTGAGCAGGTGCAGGAGAGACTGCTCAAGTCCCAAGTCCACCCAGAACTGAGGGCG 296
Oy	81 LeuMetAspGluThrrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGluLeu 100
Db	297 CTGATGACAGACCATTAAGAGATTGAAGGCTTCAACAATCGAACTGGAGAACCACTG 356
Oy	101 ThrProValAlaGluGluThrrArgAlaArgLeuserLysGluLeuGlnAlaAlaGlnAla 120
Db	357 ACCCGGGTGGGAGGAGACCGGGCAGGGCTGTCCAAAGAGCTGCAGGCGGGCCAGGCC 416
Oy	121 ArgLeuGlnAlaAspMetGluAspValCysGlnArgLeuValAGlnTrpArgGlyVal 140
Db	417 CGGCTGGCGCGGACATGGAGAGACTGTGGCGGCCCTGGTGTGCACTACCGCGCGAGGTG 476
Oy	141 GlnAlaMetLeuGlnGlnSerThrrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db	477 CAGGCGAATGCTCGGCGACAGCACCGAGAGAGTGGGGGTGGCCCTCCGCTCCACCTGGCC 536
Oy	161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTrp 180
Db	537 AAGTGGCTAAGCGGCTCTCCGCGATGCCATGACCTGCAGAGACCGCTGGCAGTGTAC 596
Oy	181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaAlaLeuArgLysArgLeuGly 200
Db	597 CAGGCGGGGGCCGCGAGGCGCGAGGCGCTGTCTACAGCGCATTCGCGACCGGCTGGGG 656

OY	201	ProleValGluIngluInglyArgValAlarValAlarThrValGlySer-LeuAlaGlyGlnPr	220
Db	657	CCCCGGTGGAGACAGGCCCGCTGGGGCCGCACATGTGGGCTCCTTGTCGGCCGACC	716
OY	220	OleuIngluArGalaglaGlnAlaTrp-GlyGuIuArg-LeuAraGlaArMetGlu--Glu	238
Db	717	GCTACAGAGCGGGGCCAGGCGCTGGGGCGCAGAGCGCTTGCGGGCGCGGATGGAGGACAT	776
OY	239	MetGlySerAlaArgThr-Arg-AspArgLeuAspGluValLys-GluInValAla---Glu	256
Db	777	TGGGACAGCCGACCCCGCGCACCGCTTGAGACAGCTAAAGAGACAGTGGCGGAAG	836
OY	257	ValArGAlaLySLeuGluGluGlu-AlaGlaIngluInlLeArgleuGlnAla--GluAlaPhe	275
Db	837	TTGGCGCCCAAGCTGGAGAGAAGCAGGCGCCAGAGATACCGCTCGACGGCCGAAGGCTTC	896
OY	276	GlnAlaArGLeuLysSerTrpPheGluProleuVal	287
Db	897	CCAGCGCGGCTCAAAGCTGTTGCAACCGCTGTT	932
RESULT 5			
LOCUS	B1600906	942 bp	mRNA linear EST 07-SEP-2001
DEFINITION	60324924F1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:5301010 5',		
ACCESSION	B1600906		
VERSION	B1600906.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
JOURNAL	1 (bases 1 to 942)		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D. Email: cgaaps-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN) CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC Clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at: http://image.llnl.gov Plate: LMNL1762 row: a column: 11 High quality sequence stop: 762.		
FEATURES			
Source	Location/Oualifiers		
	1..942		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5301010"		
	/clone_lib="NIH_MGC_96"		
	/tissue_type="hypothalamus"		
	/lab_host="DH10B"		
	/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag site)-dr primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to R0V 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."		
BASE COUNT	177 a 284 c 361 g 120 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	2.81e-117	Length:	942
Score:	1209.50	Matches:	265
Percent Similarity:	93.13%	Conservative:	6

Query Match:	Best Local Similarity:	Mismatches:	Indels:
91.07%	76.12%	8	14
DB:	13	Gaps:	1

US-09-827-854-16 (1-317) x B1600906 (1-942)

DB	Sequence	Score
OY	1 MetLysValIleuTrpAlaIleuLeuVal1ThrPheLeuAlaGlyCysGlnAlaLysVal 	20
Db	73 ATGMAAGTTCGTGGGCTCGCTTGGCTGGTACATTCTCTGGCAGATGCCAGGCCAAGGTG 	132
OY	21 GluGlnAlaValAlaGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 	40
Db	133 GAGCAAGCGGTGGAGACACAGCCGGAGCCCGAGCTCGCAGACAGACACAGTGGCAGAGC 	192
OY	41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 	60
Db	193 GGCACAGCGCTGGGAATGTGCACCTGGGTCTCTTTGGATTACTCGCTGGGTGGACACA 	252
OY	61 LeuSerGluGlnIValGlnGluLeuLeuSerSerGlnIValThrGlnGluLeuArgAla 	80
Db	253 CTGTCTGACAGAGTGCAGAGAGAGAGCTGCTCAGCTCCAGTCCACAGGAACTGAGGGG 	312
OY	81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu 	100
Db	313 CTGATGTGAGACCAACCATGAAGAGTTGTAAGAGCCCTCAAAATCGGAATCGAGGAACAATG 	372
OY	101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluGlnGlnAlaAlaGlnAla 	120
Db	373 ACCCGCGTGGCGAGAGAGAGCGGGCAGCGCTGTCCAAAGAGCTGCAGCGCGCGAGGCC 	432
OY	121 ArgLeuGlyAlaAspMetC1uAspValCysGlyArgLeuValGlnIleArgGlyGluVal 	140
Db	433 GCGGTGGCGGCGACATGGAGAGCGTGTCCGGCCGCTGTGTGACATGCCGGCGAGGTG 	492
OY	141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 	160
Db	493 CAGGCGATCTCTGGCGAGACACCGAGGAGCTCGGGGTGCGCTCCCTCCACACCTGGCC 	552
OY	161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 	180
Db	553 AAGCTGCGTAAACGGCTCTCCGCGATGCCATGACCTCAGAGAAGCGCTGGCAGGTAC 	612
OY	181 GlnAlaGlyAlaArgGluGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 	200
Db	613 CAGGCGCGGGCCCGCGAGGCGCCGAGGCGGCTCAGCCATCCGCAAGCGCTGGCGAGGTG 	672
OY	201 ProLeuValGluGlnGlyArgValArgAlaIleThrValGlySerLeuAlaGluGlnPro 	220
Db	673 CCCCTGGTGAACAGAGGCGCGGTGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 	731
OY	221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetG 	240
Db	732 CTACAGAGAGCGGCGCCAGGCTGTGGGGCGAGCGGCTGCGCGCGGATGGAGGACAAATTG 	791
OY	240 LysSerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArg 	259
Db	792 GCAGCGCGGCGCGCGCGCGCTGGAGCGGAGAGCGGAGGTTGCCGAAGTTGCCGG 	851
OY	259 AlaLysLeuGluGluGlnAlaGlnGlnGlnIleArgLeuAlaGlnAlaPheGlnAla 	278
Db	852 CCAACTGGAGGAGACACAGCGCCAGCATATACGCTTGCAGGCGCGAGGCGCTTCCAGGCC 	911
OY	278 ArgLeuLysSerTrpPheGluPro 	285
Db	912 GCCTCAAAAC-TGGTTGACCCCT 933 	

LOCUS	DEFINITION	ACCESSION	VERSION
BM042094	BM042094 800 bp mRNA linear EST 07-NOV-2001	BM042094	BM042094.1
LOCUS	60361571361 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:3420618 5'	BM042094	BM042094.1
DEFINITION	mRNA sequence.	BM042094	BM042094.1
ACCESSION	BM042094	BM042094	BM042094.1
VERSION	BM042094.1	BM042094	BM042094.1

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 800)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgarbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM1875 row: 1 column: 03
High quality sequence stop: 792.
Location/Qualifiers
1..800
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5420618"
/clone_lib="NIH_MGC_112"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOT87; Site_1: XhoI; Site_2:
ECORI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC library."
BASE COUNT 146 a 239 c 315 g 100 t
ORIGIN
Alignment Scores:
Pred. No.: 7.54e-116 Length: 800
Score: 1195.00 Matches: 243
Percent Similarity: 98.808 Conserved: 4
Best Local Similarity: 97.208 Mismatches: 3
Query Match: 75.208 Indels: 1
DB: 13 Gaps: 0
US-09-827-854-16 (1-317) x BM042094 (1-800)
QY 1 MetLeValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
DB 50 ATGAAGATTCTGTGGCTGTGCTGTGTCACATTCCTGCGCAGATGCCAGCCAGAGTG 109
QY 21 GlnGlnAlaValGlnThrGlnProGlnLeuArgGlnGlnThrGlnThrGlnThr 40
|||||
DB 110 GACCAAGCGGTGAGACAGAGCGGAGCCGAGCTCGCGCAGACCGGAGTGCGAGAGC 169
QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlnArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
DB 170 GCGCAGCGGTGAGACTGCGACTGCGCTTTGGGATTAAGCTGCGCTGGGCGAGACA 229
QY 61 LeuSerGlnGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
|||||
DB 230 CTGTCTGACAGAGTGCAGAGAGCTGCTACAGCTCCAGGTCAACGAGAACTGAGGGCG 289
QY 81 LeuMetAspGlnThrMetLeuGlnLeuLeuArgAlaTyrLysSerGlnGlnGlnGln 100
|||||
DB 290 CTGATGACAGAGCCATGAGAGGTTGAGAGGCTTACAAATCGAAGTGGAGAACTG 349
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnGlnAlaAlaGlnAla 120
|||||
DB 350 ACCCGGCTGGCGAGAGACGCGGCGACGGCTGTCCAAGAGAGTCCAGCGCGGCGAGGCC 409

QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
|||||
DB 410 CGCGTGGCGCGGACATGAGGAGCGTGTGGCGCGCTGTGTGACGACGGCGGAGGTG 469
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
DB 470 CAGGCCATGCTCGGCCAGAGACCGAGAGACTCGGGGTGGCTCCCTCCCACTGGCC 529
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
|||||
DB 530 AACCTGCTAAGCGGCTCTCCCGATGCCGATGACCTGCAGAAAGCGCTGGCAGGTGAC 589
QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
|||||
DB 590 CAGCGCGGGCGCGCAGAGCGCGCGAGCGCGGCTCAGCGCCATCCGCGAGCGCGGGG 649
QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
|||||
DB 650 CCCCTGTGGACACAGGCGCGCTGCGGCGCAGCACTGTGGCTCCTGGCGGCGCAGCGG 709
QY 221 LeuGlnGlnArgAlaGlnAlaTrpGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 240
|||||
DB 710 CTACAGAGAGCGGCGCGCGCTGGGCGAGCGGTG-CGCGCGGATGAGAGATGGCG 768
QY 241 SerArgThrArgAspArgLeuAspGlnVal 250
DB 769 ACCGGGAGCGCGAGCGCTGAGCAGCGTG 798
RESULT 7
BG472299 927 bp mRNA linear EST 21-MAR-2001
LOCUS BG472299
DEFINITION 602513830F1 NIH_MGC_16 Homo sapiens CDNA clone IMAGE:4645759 5',
mRNA sequence.
ACCESSION BG472299
VERSION BG472299.1 GI:13404485
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 927)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgarbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLM1419 row: k column: 08
High quality sequence stop: 848.
Location/Qualifiers
1..927
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4645759"
/clone_lib="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOT87; Site_1: XhoI; Site_2:
ECORI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGCAG(G). Library constructed by Ling Hong
in the Laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC library."
BASE COUNT 170 a 271 c 371 g 115 t
ORIGIN

Alignment Scores:
 Pred. No.: 1.05e-115 Length: 927
 Score: 1194.50 Matches: 274
 Percent Similarity: 92.008 Conservative: 2
 Best Local Similarity: 91.338 Mismatches: 14
 Query Match: 75.17% Indels: 12
 DB: 12 Gaps: 1

US-09-827-854-16 (1-317) x BG472299 (1-927)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||
 DB 47 ATGAAGTTCTGTGGCTCGTGTGCTGTCACATTCCTCTGCGAGATGCCAGGCCAAGGTG 106
 QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrGlnIns 40
 |||
 DB 107 GAGCAAGCGGTGGAGACAGACAGCGGAGCCGAGACTGCCAGAGACGACGAGCGAGAGC 166
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 |||
 DB 167 GGGCAGCGCTGGGAATGCGACTGGCTCTTTGGGATTACTGCGCTGGGTGCAGACA 226
 QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 |||
 DB 227 CTGTCTGACAGGTGACAGAGAGAGAGCTGCTACAGTCCAGATCACCAGAACTGAGAGCGC 286
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100
 |||
 DB 287 CTGATGAGACAGACCATGTAAGAGTTGAAAGCCTTACAAATCGGAAGTGAAGAACATG 346
 QY 101 ThrProValAlaGlnGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 |||
 DB 347 ACCCGGTGGCGAGAGAGACGCGGCGACGGCTGTCCAAAGAGCTGCAGCGCGCGAGCC 406
 QY 121 ArgLeuGlyAlaAspMetLysPylCysGlyArgLeuValGlnThrArgGlyGluVal 140
 |||
 DB 407 CGCTGGGCGCGACATGAGAGACGTGTCGCGCGCTGTGATGACGAGACCGCGAGGTG 466
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluGlnGluValArgLeuAlaSerHisLeuArg 160
 |||
 DB 467 CAGGCCATCTCGGCGAGAGACCGAGAGCTGCGGTGCGCTCCCTCCACCTGCGCC 526
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 |||
 DB 527 AAGCTCGTAAGGGCTCTCCGCGCATGCGCATGACCTCAGAGCGCTGGAGAGTAC 586
 QY 181 Gln-AlaGlyAlaArgGlyAlaGlnArgGly--LeuSerAlaLeuArgGlnArgLeu 199
 |||
 DB 587 CAGGGCGGGCGCGCGAGGCGCGCCGAGCGCCCTCAAGCGCCATCCGCGAGCGCTT 646
 QY 200 --GlyProLeuValGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyG 219
 |||
 DB 647 GGGGCGCTGGGTGGAGACAGGGCGCGGTGCGGCGCCCTGAGGCTCTGCGCGGCC 706
 QY 219 LnPProLeuGlnGlnArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGln 239
 |||
 DB 707 AGCG-CTACAGAGAGCGGGCCAGGCGCTGGGGGAAACGTG-CGCGCGGAGATGAGAGACA 764
 QY 239 eegLysSerArg-ThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlu-ValAr 258
 |||
 DB 765 TGGGGAGGGGAGCGCGAGCTGAGCCGAGGTGAAGAGAGGTGCGGGGAGGTGCG 824
 QY 258 gAlaLysLeuGlnGlnAlaGlnGlnThrLeuGlnGlnAlaGlnAla---PheGlnAl 277
 |||
 DB 825 CCGCAAGCTGGAGAGACGAGCGCAGATACGAGCTGCAGGCGGAGGGCTTCCAGGGC 884
 QY 277 aArg-LeuLysSerTrpPhe-GluProLeuValGluAspMet 290
 |||
 DB 885 CCGGCTCAAGAGAGCTGTCGAGAGCCGTGGTGAAGACATG 926
 RESULT 8
 B1597743
 LOCUS B1597743 922 bp mRNA linear EST 07-SEP-2001

DEFINITION 603248609F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5300259 5',
 mRNA sequence.
 ACCESSION B1597743
 VERSION B1597743.1 GI:15490682
 KEYWORDS Est.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiluyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM11760 row: b column: 04
 High quality sequence stop: 782.
 FEATURES
 source location/Qualifiers
 1..922
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5300259"
 /clone_1ib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: Brain; Vector: pBluescriptR (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3',
 size-selected for average insert size 2.3 kb and
 normalized to R07 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (carninci, in preparation). Library
 constructed by M. Brownstein (NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
 BASE COUNT 176 a 273 c 359 g 114 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.16e-115 Length: 922
 Score: 1191.50 Matches: 256
 Percent Similarity: 93.19% Conservative: 4
 Best Local Similarity: 91.76% Mismatches: 15
 Query Match: 74.98% Indels: 5
 DB: 13 Gaps: 1
 US-09-827-854-16 (1-317) x B1597743 (1-922)
 QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||
 DB 73 ATGAAGTTCTGTGGCTCGTGTGCTGTCACATTCCTCTGCGAGATGCCAGGCCAAGGTG 132
 QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrGlnIns 40
 |||
 DB 133 GAGCAAGCGGTGGAGACAGACGCGGAGCCGAGACTGCCAGACAGACGAGTGCAGAGC 192
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 |||
 DB 193 GGGCAGCGCTGGGAATGCGACTGGCTCTTTGGGATTACTGCGCTGGGTGCAGACA 252
 QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 |||
 DB 253 CTGTCTGACAGGTGACAGAGAGAGCTGCTACAGTCCAGATCACCAGAACTGAGAGCG 312
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
 |||

Db 313 CTGATGACAGACCATGAAGAGTTGAGGCTTACAATCGAACTGGAGACAACTG 372
Oy 101 ThpProvalAlaGlugluthrArgAlaArgLeuSerLysGluleuGlnAlaGlnAla 120
Db 373 ACCCGGTGGCGAGAGACGGCGGACGGCTGTCCAAAGAGCTGCAGGGCGGACAGCC 432
Oy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
Db 433 CGGCTGGCGCGGACATGAGGAGCGTGTGGCGCGCTGTGCATGACCCGGCGAGAGTG 492
Oy 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 493 CAGGCCATGCTCGGCCACAGACACCGAGAGCTCGGGTGGCGCTCCGCTCCACCTGGCC 552
Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTrp 180
Db 553 AAGCTGGTAAAGCGGCTCTCCGCGATGGCGATGACCTGCAGAAAGCGGCTGGCAGGTGTAC 612
Oy 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 613 CAGGCGGGGCGCGCAGAGGCGCGGAGCGCGCTC-AGCGCCATCGCGGAGCGCGCGGGG 671
Oy 201 ProLeuVal-GluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGln-P 220
Db 672 CCCCTGGTGGGAACAGAGCGCGGGTGGCGCGCGCACTGTGGCTCCCTGGCGGCGCAAGC 731
Oy 220 roLeuGlnGluArg-AlaGlnAlaTrpGlyGlyArgLeuArgAlaArgMetGluGlnMet 239
Db 732 CGCTACAGAGAGCGGCGCGCGCTGGGCGAGCGGATACGCGCGGATGAGCAGCAGCAG 791
Oy 240 GlySerArgThrArgAspArgLeuAspGlyValLysGlnGlnValAlaGluValArgAla 259
Db 792 GGGAGCGGCGGACCGGACCGCTGAGCAGAGTGAAGAGAGAGCGGCGGGGCGGCGCA 851
Oy 260 LysLeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGln 276
Db 852 CAAGCTGAAGAAACAGCCCAAGAAATACGCTTA--GCCGAAAGCTTTCAGG 899

RESULT 9
Bg774871 817 bp mRNA linear EST 15-MAY-2001
LOCUS Bg774871
DEFINITION 602649975F1 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:4760827 5',
mRNA sequence.
ACCESSION Bg774871
VERSION Bg774871.1 GI:14045188
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 817)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LRCM1612 row: e column: 20
High quality sequence stop: 813.
Location/Qualifiers
1..817
/Organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4760827"
/clone_lib="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"

/note="Organ: prostate; Vector: pORF7; Site_1: XhoI;
Site_2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI using the
following 5' adaptor: GCCACAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC library."

BASE COUNT 148 a 242 c 324 g 103 t
ORIGIN

Alignment Scores:
Pred. No.: 1,62e-114 Length: 817
Score: 1182.50 Matches: 251
Percent Similarity: 96.93% Conservative: 2
Best Local Similarity: 96.17% Mismatches: 6
Query Match: 74.42% Indels: 4
DB: 12 Gaps: 0

US-09-827-854-16 (1-317) x Bg774871 (1-817)

Oy 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 40 ATGAAGTTCTGTGGGCTGCTGCTGTGTCACATTCCTGGCAGAGATCCAGGCGCAAGTG 99
Oy 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 100 GAGCAAGCGGTGGAGACAGAGCGGAGCGCGAGCTCGCGCAGAGACCAAGTGGCAGAC 159
Oy 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
Db 160 GGCACGCGCTGGGAACCTGGACACTGGGTCTTTGGATTACCTGGCTGGGGTGGAGCA 219
Oy 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 220 CTGCTGACAGAGTGCAGAGAGAGCTGCTCAGCTCCAGTACCCAGAGACTGAGCGGG 279
Oy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu 100
Db 280 CTGATGAGCAGACCATGAAGAGTGAAGGCTTACAATTCGAACTGAGAGCAACTG 339
Oy 101 ThpProvalAlaGlugluthrArgAlaArgLeuSerLysGlnGlnAlaGlnAla 120
Db 340 ACCCGGTGGCGAGAGAGCGGCGGCGGCTGTCCAAAGAGCTGCAGGCGGCGAGGCC 399
Oy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
Db 400 CGGCTGGCGCGGACATGAGAGACGTGTGCGCGCTGTGTCAGTACCGCGGCGAGGTG 459
Oy 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 460 CAGGCCATGCTCGGCGCAGAGACCGGAGAGCTGCGGCTCGCGCTCCGCCACCTGGCC 519
Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTrp 180
Db 520 AAGCTGCGTAAAGCGGCTCTCCGCGATGCGGATCTGCAGAAAGCGCTGGCAAGGTAC 579
Oy 181 GlnAlaGlyAlaArgGluGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 580 CAGGCGGGGCGCGGAGAGGCGCGGCGGCTGAGGCCATCCGCGAGGCGGCTGGGG 639
Oy 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 640 CCCCTGGTGAAGAGAGGCGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 698
Oy 221 LeuGlnGluArgAlaGlnAla--TrpGlyGluArgLeuArgAlaArgMetGluGlnMetG 240
Db 699 CTACAGAGAGCGGCGCGGCTTGGGCGGAGGATGCGCGCGGATGGAGAGATGG 758
Oy 240 LysSerArgThrArgAspArgLeuAspGlyValLysGlnGlnValAlaGluValArgAla 259
Db 759 GCCACCGGAGCG--GAACCGCTGGAGAGAGTGAAGAGCAGAGGTGGCGGAGGTGGCGGCC 815


```

RESULT 10
BI600563
LOCUS      603244936F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5287329 5',
DEFINITION mRNA sequence.
ACCESSION BI600563
VERSION    BI600563.1 GI:15493502
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 811)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
AUTHORS    Contact: Robert Strausberg, Ph.D.
            Email: cga@bbs-remail.nih.gov
            Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNLN at:
            http://image.llnl.gov
            Plate: L1AM1726 row: g column: 10
            High quality sequence stop: 783.
FEATURES
            source
            1..811
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5287329"
            /clone_lib="NIH_MGC_96"
            /tissue_type="hypothalamus"
            /lab_host="DH10B"
            /note="Organ: brain; Vector: pBluescriptR (modified
            pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
            ); Oligo-dT primed using primer 5'-TTTTTATTTTATTTTATTTT-3',
            size-selected for average insert size 2.3 kb and
            normalized to 10^6 5. This is a primary library enriched
            for full-length clones and constructed using the
            Cap-trapper method (Carninci, in preparation). Library
            constructed by M. Brownstein (NIH/NHGRI, National
            Institutes of Health). Note: this is a NIH_MGC library."
BASE COUNT      148 a      245 c      316 g      102 t
ORIGIN
Alignment Scores:
Pred. No.:      2,31e-114      Length:      811
Score:          1181.00      Matches:      240
Percent Similarity: 97.97%      Conservative: 1
Best Local Similarity: 97.56%      Mismatches: 4
Query Match:    74.32%      Indels:      1
DB:             13      Gaps:      0
US-09-827-854-16 (1-317) x BI600563 (1-811)
OY      1 MettysValleuTPtAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db      74 ATGAAGGTTCTGTGGTGGCTGGTTCATCTCTGCGAGATGCCAGGCCAAGCTG 133
OY      21 GUGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db      134 GAGCAAGGCGGTGGAGACAGAGCCGAGCTGCGCCACAGACAGCGAGGCGAGAGC 193
OY      41 GAGlnATGTPGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTyrValGlnThr 60
Db      194 GCGCAGCGCTGGAGACTGGCACTGGCTTGTGGATTAACCTCGCTGGGTGCAGACA 253
OY      61 LeuSerGluGlnValGlnGluGlnLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db      254 CTGTCTGAGCAGAGTGCAGAGAGCTCTCAGCTCCAGAGTCCAGAGAACTGAGGCG 313

```

```

OY      81 LeuMetAspGluThrMetLeuGluLeuLysAlaTyrLysSerGluLeuGluGluGln 100
Db      314 CTGATGAGACGAGACCATGATGAAGCTTGAAGCTTCAATGAGGAACTGGAGAACTG 373
OY      101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db      374 ACCCGGTGGCGGAGAGACACGGGGCGACGGCTGTCCAGAGAGTCCAGCGCGGAGGCG 433
OY      121 ArgLeuGlyAlaAspMetGluAspValLysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db      434 CGGCTGGGCGCGGACATGAGAGACGTGTGCGGCCCTCTGTGCATGCTACCGCGGAGGCTG 493
OY      141 GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db      494 CAGCGCATGCTGGCCCAAGACACGAGAGAGCTGGCGGGTGGCTTGCCTCCACTGGCGC 553
OY      161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db      554 AAGCTGCGTAAAGCGGCTCTCCGCGATGCGATGACAGAAAGCGCTGGCACTGATAC 613
OY      181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeu-Gl 200
Db      614 CAGCGCGGGCGCGCGAGGCGCGCAGCGCGCTCAGCGCCATCCGAGAGCGCTGGGG 673
OY      200 YProLeuValGluGlnArgValArgValAlaAlaThrValGlySerLeuAlaGlyGlnPr 220
Db      674 CCCCTGTGTGAGACAGGGCCCGCTGCGGGCCGCCACTGTGGGCTCTCGCCAGGACACC 733
OY      220 OLeuGlnGluArgAlaGlnAlaTatPGLyGluArgLeuArgAlaArgMetGluGluMetGl 240
Db      734 GCTACAGAGAGCGGGCCAGGCGCTGGGGCGAGGCTGGCGGGCGGAGTGAAGCAATGG 793
OY      240 YSerArgThrArgAsp 245
Db      794 CAGCGGAGACCGCGAA 809
RESULT 11
BI6029472
LOCUS      602763768F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4899112 5',
DEFINITION mRNA sequence.
ACCESSION BI6029472
VERSION    BI6029472.1 GI:4177059
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 845)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
AUTHORS    Contact: Robert Strausberg, Ph.D.
            Email: cga@bbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNLN at:
            http://image.llnl.gov
            Plate: L1CM1790 row: g column: 17
            High quality sequence stop: 829.
FEATURES
            source
            1..845
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4899112"
            /clone_lib="NIH_MGC_42"
            /tissue_type="epithelioid carcinoma cell line"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;

```


	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
OY	61	LeusertGlInglValGlnGluGluIleuLeuSerSerGlnValThrGlnGluLeuAArgAla	80		
Dd	253	CtGTCTGCACAGTGTACAGAGAGCGTGCTCAGCTCCAGTCACCAGAACGTGAGGGCG	312		
OY	81	IeuMetAspGIuThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluIleuLeu	100		
Dd	313	CtGATGACAGACCATTAAGAAGATTGAAGCCCTCAAAATGTGGACTGAGACACAATCG	372		
OY	101	ThrProValAlaGluGluIuThrArGlaAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120		
Dd	373	ACCCGGTGGCGGAGAGACGGCGGCCACAGGGGTGTCCAAGAGAGCTCAGCGCGCACGGCC	432		
OY	121	ArgLeuGluAlaAspMetGluAspValLysGluArgLeuValGlnTrpArgGlyGluVal	140		
Dd	433	CGGCTGGGCGCGACATGAGAGACTGTGTGGCCGCCCTGTGTGCATACCGCGCGAGGTG	492		
OY	141	GlnAlaMetLeuGlyGlnSerThrGlnGluLeuLeuArgValAlaArgLeuAlaSerHisLeuArg	160		
Dd	493	CAGGCAATGCTCTGGCCAGACACCGAGAGAGCTGCGGGTGGCTGTCCCTCCACTGGCG	552		
OY	161	LysLeuArgLysArgGluLeuLeuArGspAlaAspAspLeuGlnLysCysLeuAlaValTyr	180		
Dd	553	AAGCTGGCTAAGCGGGCTCTCTCCGCATGCCATGACCTGCAGAAACGGCTGGCACTGTAC	612		
OY	181	GlnAlaGlyAlaArgGluGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200		
Dd	613	CAGCCCGGGGGCCCCGAGAGCGCCGAGCGCGGTCTCACACGCCATCCGCGAGCGCTGGGG	672		
OY	201	ProLeuValGlnGlnGlyArgValAlaArgAlaAlaTrpValGlySerLeuAlaGlyLntro	220		
Dd	673	CCCCTGGTGGAACACAGGCGCGCTGGGGCCGCCACTGTGGGTCTCTCTGGCCACGACCGCG	732		
OY	221	LeuGlnGluArGalaGlnAlaAlaTrpLysGlyLnuArgLeuArGalaArgMetGluGluMet	239		
Dd	733	CTACAGAGAGCGGGCCAGGCTCGGGGGCCAGCCGCTG-CGCGCGCGATGAGAGAGATG	788		
RESULT 13	BIS51475	60319431FPI NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5274003 5',			
LOCUS	BIS51475	mRNA sequence.			
DEFINITION	BIS51475	BIS51475.1 GI:15438787			
ACCESSION	BIS51475	EST.			
VERSION	BIS51475.1	human.			
KEYWORDS		Homo sapiens			
SOURCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
ORGANISM		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
		1 (bases 1 to 919)			
		NIH-MGC http://mgs.nci.nih.gov/.			
		National Institutes of Health, Mammalian Gene Collection (MGC)			
		Unpublished (1999)			
		Contact: Robert Strausberg, Ph.D.			
		Email: cgapbs-remail.nih.gov			
		Tissue Procurement: Miklos Palkovits, M.D., Ph.D.			
		cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shizaki			
		Toshiyuki and Piero Carninci (RIKEN)			
		cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)			
		DNA Sequencing by: Incyte Genomics, Inc.			
		Clone distribution: MGC clone distribution information can be			
		found through the I.M.A.G.E. Consortium/LMU at:			
		http://image.llnl.gov			
		Plate: LAM11691 row: 1 column: 04			
		High quality sequence stop: 812.			
FEATURES		Location/Qualifiers			
SOURCE		1..919			
		/organism="Homo sapiens"			
		/db_xref="taxon:9606"			
		/clone="IMAGE:5274003"			
		/clone_lib="NIH_MGC_95"			
		/tissue_type="hippocampus"			
		/lab_host="DH10B"			

```

/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgagaa
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.5 kb and
normalized to 10^6 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      172 a      270 c      363 g      114 t
ORIGIN
Alignment Scores:
Pred. No.:      4,58e-110      Length:      919
Score:          1141.00      Matches:      246
Percent Similarity: 93.26%      Conservative: 3
Best Local Similarity: 92.13%      Mismatches: 17
Query Match:      71.61%      Indels:      4
DB:              13      Gaps:      0
US-09-827-854-16 (1-317) x B1551475 (1-919)

```

QY	1	MellysValLeuTPRPaLaLaLeuLeuValThrpheLeuAlaGlyCysGlnAlaLysVal	20
Db	73	ATGAAGGTTCTGTGGGCTGGCTGTGCTGCATCTCCGGCAGATGCCAGGCCAAGGTG	13
QY	21	GIuGlnAlaValaGIuThhGIuProGIuProGIuLeuAArgGlnInThhGIuTrpGlnSer	40
Db	133	GAGCAAGCGGTGGACACAGAGCCGGAGCCCGAGCTGCCCCACAGACCAGATTGGCAGAGC	19
QY	41	glycInArGTpRGluLeuAlaLeuGlyArpRhpheTrpAspTyrLeuArGTpValaGlnThr	60
Db	193	GGCCAGCGCTGGAACTGGCACTGGGTGCTTTGGAGTTACTGTGGCTGGGTGGAGACA	25
QY	61	LeuSerGIuGlnValaGlnGIuLeuLeuSerSerGlnValThhGlnIuLeuArGla	80
Db	253	CTGTCTGAGCAGGTGGACAGAGAGAGCTGCTCACTCCAGGTACCACCAAGACTGAGGGCG	31
QY	81	LeuMetAspGIuThrMelLysGIuLeuLysAlaTyrLysSerGIuLeuGlnGIuGlnLeu	100
Db	313	CTGATGGACGAGACCATAGAAGATTGAAGGCTTACAATGTGAGACTGGAGAACAACTG	37
QY	101	ThrpProValaGIuGIuGIuThrArGlaArArgLeuSerLysGIuLeuGlnAlaGlnAla	120
Db	373	ACCCGGGTGGGAGAGAGACGGGGCCACGGGTGTCACAGAGAGCTGCAAGCGGAGGCC	43
QY	121	ArgLeuGlnAlaAspMetGIuAspValaIcysGlyArGleuValaGlnTyrArGlyGIuVal	140
Db	433	CGGCTGGGCGCGACATGAGAGACTGTGGCGGCCCTGTGGTAGGTACCGCGGAGGTG	49
QY	141	GlnAlaMetLeuGlyGlnSerThhGIuGIuLeuArGValaArgLeuAlaSerHisLeuArg	160
Db	493	CAGGCATGCTGGCGCAAGACCGAGGAGTGGGGGTGCTCGCTCCACCTGGCC	55
QY	161	LysLeuArGlyAsArgLeuLeuArGAspAlaAspAspleuGlnLysCysLeuAlaValTyr	180
Db	553	AAGCTGCTTAAGGGCTCTCTCCGCATCCGCATGACCTGCAGAAAGCGCTGGCAGTGTAC	61
QY	181	GlnAlaGIuAlaArGlyGIuAlaGlnArGlyLeuSerAlaIaIeArgGIuArGleuGly	200
Db	613	CAGCGCGGGGGCGGAGTGGCGGAGCGGCTGTCTACAGCCGCAATCCGAGACGCGCTGGG	67
QY	201	ProLeuValaGIuGlnGlyArGValaArgAlaAlaThrValaGlySerLeuAlaGlyGlnPro	220
Db	673	CCCCGGTGGGAACAGGCGCGGTGGGGCCGCCACTGGGTGGCTCCGTGGCGGAGCCCT	73
QY	221	LeuGlnGIuArGlaGlnAlaIATpGIuArGleuArGlaArGmetGIuGIu-MetGI	240
Db	733	A--CAGAGACGGGGCGACAGCCTGGGGCGAGCGGTGGCGGCCCGGATGGAGAAATGTGG	79
QY	240	ySerArGTpThrArGAspArGleuAspGIuValaLysGIuGlnValaGlnValaArgAlaLys	260
Db	791	CAGCGGACCGCGGGACG--CTGGACGAAGTGAAGAGCACTGGGCGGGAGGG--TCCGCCAA	84

QY 260 sleugluglunlaagln 266
|||||:::|
Db 848 GCTGAGAGACACAGAG 866

RESULT 14
BM042228

LOCUS BM042228 757 bp mRNA linear EST 07-NOV-2001
DEFINITION 603616186F1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5557004 5',
mRNA sequence.
ACCESSION BM042228
VERSION BM042228.1 GI:16771495
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 757)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTP/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2021 row: k column: 21
High quality sequence stop: 757.
Location/Qualifiers

FEATURES
source

1..757
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5557004"
/clone_lib="NIH_MGC_112"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site: 1: XhoI; Site: 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."

BASE COUNT 134 a 229 c 236 g 98 t

ORIGIN

Alignment Scores:
Pred. No.: 3.95e-109 Length: 757
Score: 1131.00 Matches: 234
Percent Similarity: 98.738 Conservative: 0
Best Local Similarity: 98.738 Mismatches: 2
Query Match: 71.188 Indels: 2
Gaps: 0

US-09-827-854-16 (1-317) x BM042228 (1-757)

QY 1 MettysValleuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
Db 50 ATGAAAGTCTGTGGCTCGTGTGTCACATTCCTGCGAGAGATGCCAGGCCAGGTG 109
|||||

QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGluThrGluTrpGlnSer 40
|||||

Db 110 GAGCAAGCGGTGAGACAGAGCCGAGAGCCGAGCTCGCCAGAGACGAGGTGCAAGC 169
|||||

QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||

Db 170 GGCAGACGCTGGAACTGGCAGCTGGCTCTTTGGGATTACTGGCGCTGGGTGCAGACA 229
|||||

QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
Db 230 CTGTGACGACAGCTGCAGAGAGCTCAGCTCCAGCTCACCCAGAACAGAGAGCGG 289
|||||

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
|||||

Db 290 CTGTATGACAGACATGAAAGAGTTGAAGCTTACAAATTCGAATCGAGACAACTAG 349
|||||

QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
|||||

Db 350 ACCCGGCTGGCGAGAGACAGCGGGCAGCGCTGTCCAAAGAGCTGCAGCGCGGCGG 409
|||||

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
|||||

Db 410 CGGCTGGGGCGGACATGAGAGACGTGTGCGGCGCTGTGAGTACCGCGCGGAGTG 469
|||||

QY 141 GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||

Db 470 CAGGCCATGCTCGCCAGAGACCGAGAGCTCGGGGCTCGCTCCACCTGCGC 529
|||||

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
|||||

Db 530 AAGCTGCTGAAGGGGCTCTCTCGCGCATGCGCATGACCTGCAGAGCGCCTGGCAGTAC 589
|||||

QY 181 GlnAlaGlyAlaArgGluGluGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||

Db 590 CAGCGCGGGGCGCCAGAGGCGCGCGCGCTCAGCGCATCCGCGAGCGCGCTGGG 649
|||||

QY 201 ProLeuValGluGluGlnGlyArgVal-ArgAlaAlaThrValGlySerLeuAlaGlyGln 220
|||||

Db 650 CCCCTGCTGGACAGAGGCGCGCTGGCGCGCGCTGAGGCTCCCTGCGCGCGCAGCC 709
|||||

QY 220 OleuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMet 236
|||||

Db 710 GCTACAGAGAGGCGC-CAGGCTGGGGGAGCGGCTGCCGCGCGGATG 757
|||||

RESULT 15
BG706129

LOCUS BG706129 907 bp mRNA linear EST 07-MAY-2001
DEFINITION 602669093F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4792030 5',
mRNA sequence.
ACCESSION BG706129
VERSION BG706129.1 GI:13981169
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 907)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM10669 row: 1 column: 23
High quality sequence stop: 832.
Location/Qualifiers

FEATURES
source

1..907
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4792030"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified

Bluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTN-3',
size-selected for average insert size 2.3 kb and
normalized to R0.5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 171 a 260 c 358 g 117 t 1 others

ALIGNMENT SCORES:

Pred. No.:	5.14e-108	Length:	907
Score:	1121.50	Matches:	243
Percent Similarity:	94.23%	Conservative:	2
Best Local Similarity:	93.46%	Mismatches:	11
Query Match:	70.58%	Indels:	6
DB:	12	Gaps:	2

US-09-827-854-16 (1-317) x BG706129 (1-907)

```
OY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 75 ATGAAGGTTCTGTGGCTGCTGTTGCTGCATTCTGCGAGATGCCAGGCAAGGTG 134
OY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 135 GAGCAAGCGGTGAGACAGACAGCGCGGAGCCGAGCTGCCGACAGACCCGAGTGGCAGAGC 194
OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 195 GGCCAGCGCTGGAGACTGAGCTGCTGCTTTGGATTACCTGCGCTGGGTGGCAGACA 254
OY 61 LeuSerGluGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 255 CTGTGTAGACAGTGCAGAGAGAGCTGCTCAGCTCCAGGTCACCCAGAACTGAGGCG 314
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
DB 315 CTGATGAGACAGACCATGATGAGAGTTGAAAGCCTTACAAATCGAATCGAGAGACACACTG 374
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGlu-LeuGlnAlaAlaGlnAl 120
DB 375 ACCCGGTGGCGGAGAGACGCGGCGACGCTGTCCAAAGAACTGCAAGCGCGCGCAGGC 434
OY 120 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 435 CCGGCTGGCGCGGACATGAGAGACGTGTGCGCGCCCTGTGTGACGTACCGCGCGAGGT 494
OY 140 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 495 GCGAGCCATGCTCGCGCCAGAGACCCGAGAGAGTGGCGGTGCCCTGTGCTCCACACTGCG 554
OY 160 GlnLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
DB 555 CAAGCTGCGTAAAGCGGCTCTCCGCGATGCCATGACCTGCAGAAAGCGCTGGCAGTGA 614
OY 180 GlnAlaGlyAlaArgGlnGlnGlyAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeu-G 200
DB 615 CCAAGCGCGGCGCGCGAGGCGCGGAGCGGCGCTCAAGCGCCATCCGAGCGCGCTGGG 674
OY 200 LysProLeuValGlnGlnGlnLysArgValArgAlaAlaThrValGlnSerLeuAlaGlnL 220
DB 675 GCCCTTGTGTGAACAGGCGCGCGTGGCGCGCCACTGTGGGCTCCTGAGCGGCGCAGC 734
OY 220 LeuGlnGlnLysArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnLysMetG 240
DB 735 CGCTACAGAGAGCGCGCGAGGTG--GGGCGAGCGGTGCGCGCGGATGAGAGAGATGG 791
OY 240 LysArgTrpArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArg 258
DB 792 GCAG-CGAGCCCGG--ACGCTGACGAGTG-AAAGAGCAGTGGGGAAGTGTGCGC 842
```

Search completed: March 14, 2003, 20:14:03
Job time : 1293.69 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:08:57 ; Search time 2307.23 Seconds
(without alignments)
3998.551 Million cell updates/sec

Title: US-09-827-854-17
Perfect score: 1589
Sequence: 1 MKVLMALLVFLAGCOAKV.....VEKVOAVGTSAPVPSDNH 317

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlp
-O=/sgn2.1/USPRO/spool/US09827854/rumat_11032003_101610_27486/app.query.fasta.1.3576
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854 @CGN.1.13008 @rumat_11032003_101610_27486 -NCPU=6 -ICPU=3
-NO_XLPYX -NO_MMAP -LARGQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pt:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_com:*
21: em_cor:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrl:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1582	99.6	1156	6	BD004277 Apo E hum
2	1570	98.8	1110	6	E00359 E00359 cDNA coding
3	1570	98.8	1110	6	E00823 E00823 cDNA sequence
4	1570	98.8	1147	6	AX302545 AX302545 Sequence
5	1570	98.8	1156	6	BD004278 Apo E hum
6	1570	98.8	1156	9	BD004278 Apo E hum
7	1570	98.8	1186	9	K00396 Homo sapien
8	1563	98.4	1110	6	BC003557 BC003557 Homo sapi
9	1561	98.2	1157	6	E08423 E08423 DNA coding
10	1561	98.2	1157	6	AX333278 AX333278 Sequence
11	1561	98.2	1157	6	AX409597 AX409597 Sequence
12	1561	98.2	1157	6	HA04957 HA04957 Sequence
13	1558	98.0	1156	6	BD004279 BD004279 Apo E hum
14	1436	90.4	1178	9	MEAPOE MEAPOE
15	1381.5	86.9	5491	9	AF261272 AF261272 Sequence
16	1381.5	86.9	41907	6	AX358722 AX358722 Sequence
17	1381.5	86.9	107567	9	AF050154 AF050154 Homo sapi
18	1381.5	86.9	107567	9	AC011481 AC011481 Homo sapi
19	1364	85.8	5515	9	HM0460E4 HM0460E4
20	1335	84.0	5413	9	AF261280 AF261280 Pan trogl
21	1252	78.8	4762	9	BABAPOE BABAPOE
22	1251	78.7	208239	2	AC021988 AC021988 Homo sapi
23	1243	78.2	1138	4	AF303830 AF303830 Tupia gl
24	1155	72.7	1060	4	RABAPOLP RABAPOLP
25	1148.5	72.3	965	6	AX384545 AX384545 Sequence
26	1148.5	72.3	1108	4	BTAPOLPE BTAPOLPE
27	1148.5	72.3	5617	6	AX384541 AX384541 Sequence
28	1148.5	72.3	6026	6	AX384539 AX384539 Sequence
29	1136	71.5	718	9	AF200497 AF200497 Pan trogl
30	1134	71.4	1045	10	MUSAPOE MUSAPOE
31	1134	71.4	1104	10	BC028816 BC028816 Mus muscu
32	1132	71.2	718	9	AF200500 AF200500 Gori11a g
33	1132	71.2	718	9	AF200503 AF200503 Pongo pyg
34	1130.5	71.1	1154	4	BTAPOLPE BTAPOLPE
35	1128	71.0	718	9	AF200506 AF200506 Hylobates
36	1128	71.0	1122	4	SSAPOE SSAPOE
37	1123	70.7	1126	6	ARI64342 ARI64342 Sequence
38	1123	70.7	1126	6	AR205885 AR205885 Sequence
39	1079	67.9	951	10	S76779 S76779 RAPOE-apol1
40	1057	66.5	959	10	MUSAPOE MUSAPOE
41	1041.5	65.5	1069	10	RATAPOE RATAPOE
42	1014.5	63.8	228698	2	AC127479 AC127479 Mus muscu
43	1014	63.8	237653	2	AC073760 AC073760 Mus muscu
44	988	62.2	4856	10	MUSAPE MUSAPE
45	987	62.1	4267	6	ARI64387 ARI64387 Sequence

RESULT 1

ALIGNMENTS

	BD004277	BD004277	1156 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	Apo E humanized mammal.					
DEFINITION	Bd004277					
ACCESSION	BD004277.1	GI:18632238				
VERSION	JP 2001017028-A/1.					
KEYWORDS	Homo sapiens.					
SOURCE	Homo sapiens.					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE	1 (bases 1 to 1156)					
AUTHORS	Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.					
TITLE	Apo E humanized mammal					
JOURNAL	Patent: JP 2001017028-A 1 23-JAN-2001;					
COMMENT	MITSUBISHI CHEMICAL CORP OS Homo sapiens (human) PN JP 2001017028-A/1 PD 23-JAN-2001 PF 28-APR-2001 JP 2000128919 PR					
FEATURES	FT SHINOBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA PC A01667/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10, PC C12N15/09//C07K14/775, PC (C12M5/10,C12R1:91), C12N5/00,C12N15/00,(C12N5/00,C12R1:91) CC	Key Location/Qualifiers (61),(1011).				
BASE COUNT	208 a 367 c 432 g 149 t					
ORIGIN						
Alignment Scores:						
Pred. No.:	1,59e-84	Length:	1156			
Score:	1582.00	Matches:	316			
Percent Similarity:	99.68%	Conservative:	0			
Best Local Similarity:	99.68%	Mismatches:	1			
Query Match:	99.56%	Indels:	0			
DB:	6	Gaps:	0			
US-09-827-854-17 (1-317) x BD004277 (1-1156)						
OY 1	MetlyValLeutrpAlaAlaleuValThrPhelenuAGlyCysGlnAlaLysVal					
Db 61	ATGAAGCTTCGTGGCGTGCCTTGCTGCATCTCCGGCAGATGCCAGGCCAAGTG					
OY 21	GlUGlAlaValAGluThcLupProGLuPProGLuLeuArgGlnInThcLupTrpInsSer					
Db 121	GAGCAAGCGGTGGAGAACAGAGCCGAGCCAGCTGCSCCAGACAGACCAGTAGGCAGAC					
OY 41	GLyGlnArGYTPgLUleuAlaLeuGLylarphetrapPyrrleuArgTrpValGlnThr					
Db 181	GGCCAGCCCTGGGAACCTGGCACTGGGTGCTTTTGGATTACTCTCGCTGGGGTCAGACA					
OY 61	LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrnGlnGluLeuArgLa					
Db 241	CTGTCTGAGCAGGTGCACAGAGAGCTGCTCACGTCCAGGTCCACCCACAGAACTAGAGSGC					
OY 81	LeuMetaspGluThrMetLysGluLeuLysAlaTryLysSerGluLeuGluGluGlnLeu					
Db 301	CTGATGAGCAGACCATTAAGAAGATTGAAGCCCTTCACAAATCGAAGCTGAGAACAACTG					
OY 101	ThrProVALAGluGluThrFARGAlaArgLeuSerLysGluLeuGlnAlaAGlnAla					
Db 361	ACCCGGGTGGGAGGAGAACCGGGCAGCGCTGTCCAAGAGCTCAGCGCGCCAGAGCC					
OY 121	ArgLeuGluAlaAspMetGluAspAlaCysGLyArgLeuValGlnTrpArgGluVal					
Db 421	CGGCTGGGGCGGAGCATGTGAGAGACGTGTGGCGCCCTCGTGTGCAGTACCGCGGCGAGGTG					
OY 141	GlnAlaMetLeuAspGlnSerThrcLugluLeuArgValArgLeuAlaSerHisLeuArg					

DB	LOCUS	DEFINITION	CDNA coding	human	apoliipoprotein E3.	RNA	linear	PAT	29-SEP-1997															
Db	481	CAGGCCATCTCTGGCCACAGACACGAGAGCTGCGGGTCCGCTCCACCTCGCC							540															
Qy	161	LysLeuAArgLysArgLeuLeuAArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr							180															
Db	541	AAGCTGCGTAACCGGCTCTCTCCGCGATGCGGATGACCTCACAAGAGCGCTGGCAGGTAC							600															
Qy	181	GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln							200															
Db	601	CAGCGCGGGGGCCGCGAGGGCCGCGAGCGGGCTTCAGCCCATCCGCGAGCGCTGGGG							660															
Qy	201	ProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln							220															
Db	661	CCCCTGCTGGTGAACAGAGGCGCGCTGCGGGCGCCGACCTGTGGCTCCCTGGCCGAGCCG							720															
Qy	221	LeuGln							240															
Db	721	CTACAGAGACGGGGCCGAGCGCTGCGGGCGAGCGGCTGCGGGCGCGATGAGAGATGGGC							780															
Qy	241	SerArgThrArgAspArgLeuAspGlnValLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGln							260															
Db	781	AACCGGACCGCGGAGCGCTCTGGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG							840															
Qy	261	LeuGln							280															
Db	841	CTGGAGAGACAGGCGCCAGCAGATACGCTCTGACAGCGCGAGGCTTCCAGGCGCGCTCAG							900															
Qy	281	SerTyrPheGlnProLeuValGlnLysPheMetGlnArgGlnTyrPheGlnGlnGlnGlnGln							300															
Db	901	ACCTGGTTCGACCCCTCTGCTGGAGACATGACGAGCGCATGTGGCGGCTGTGTGAGAAAG							960															
Qy	301	ValGlnAlaAlaValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln							317															
Db	961	GTCAGGCTGCGCTGGGACACGAGCGCGCGCTGTGCGCCAGCACAATCAC							1011															
RESULT 2																								
E00359	E00359	1110 bp	RNA	linear	PAT	29-SEP-1997																		
LOCUS	DEFINITION	CDNA coding	human	apoliipoprotein E3.																				
ACCESSION	E00359																							
VERSION	E00359.1	GI:2168646																						
KEYWORDS	JP 1985118189-A/1.																							
SOURCE	Homo sapiens.																							
ORGANISM	Homo sapiens.																							
REFERENCE	Eukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;																							
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.																							
TITLE	Tetranishi, Y., Takamatsu, N., Matsui, Y., Kimura, M. and Ikeda, Y.																							
JOURNAL	DNA FRAGMENT																							
	Patent: JP 1985118189-A 1 25-JUN-1985;																							
	MITSUBISHI CHEM IND LTD																							
COMMENT	OS human																							
	PN JP 1985118189-A/1																							
	PD 25-JUN-1985																							
	PE 29-NOV-1983 JP 1983224980																							
	PI TERANISHI YUTAKA, TAKAMATSU NOBUHIKO, MATSUI YASUSHI, PI																							
	KIMURA MASAKO,																							
	PI IKEDA YASUOKO																							
	PC C12N15/00, C07H21/04//C12P21/00;																							
	CC strandedness: Double;																							
	CC topology: Linear;																							
	CC hypothetical: No;																							

source	1. 1110	/organism="Homo sapiens"			
BASE COUNT	198 a	353 c	416 g	143 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	7.67e-84	Length:	1110		
Score:	1570.00	Matches:	315		
Percent Similarity:	99.37%	Conservative:	0		
Best Local Similarity:	99.37%	Mismatches:	2		
Query Match:	98.80%	Indels:	0		
DB:	6	Gaps:	0		
US-09-827-854-17 (1-317) x E00359 (1-1110)					
QY	1	MethylsValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal	20		
DB	15	ATGAAGGTTCTGTGGGCTGTGCTGTGCATTCCTGCGAGATGCCAGGCCAAGGTG	74		
QY	21	GIuGlnAlaValaGluThrGluProGluProGluLeuArgGlnGlnThrGlnTrpGlnSer	40		
DB	75	GAGCAACGGGTGGAGACAGAGCCGGAGCCGAGCTGCGCAGAGAGCCAGTGGCAGAC	134		
QY	41	GIyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr	60		
DB	135	GGCCAGCGCTGGGAACAGCGACCTGGCTTTTGGATTACCTGGCGGTGGCAGACA	194		
QY	61	LeuSerGluGlnValaGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80		
DB	195	CTGTCTAGCAGGTGCGAGAGAGAGCTGTCTACGCTCCAGAGTCAACCAGAGACTGAGGGCG	254		
QY	81	LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu	100		
DB	255	CTGATGGAGACAGACCATGAGAGAGTTGAAAGGCTTACAAATCCGAAGTGGAGACACTG	314		
QY	101	ThrProValaIaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaGlnAla	120		
DB	315	ACCCGCGTGGCGAGAGAGCGGGCGCGCTGTCCAAAGAGCTGACAGCGCGCGCAGCGC	374		
QY	121	ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal	140		
DB	375	CGGCTGGGGCGGACATGAGAGAGAGCTGTGCGCGCGCTGTGCAGTACCGCGCGAGGTG	434		
QY	141	GlnAlaMetLeuAspGlnSerThrGluGluLeuArgValaArgLeuAlaSerHisLeuArg	160		
DB	435	CAGGCCATCTCTGGCCAGAGACACCGAGAGCTGGGGTGGCGCTGCTCCCACTGGCGC	494		
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr	180		
DB	495	AGGCTGCGTAAGCGGGTCTCTCGCGATGCGGATGACCTCAGAAAGCGCTGGCAGGTAC	554		
QY	181	GlnAlaGlyAlaArgGluGluGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200		
DB	555	CAGGCCGGGGCGCGAGGGCGCGAGCGGGCTCAGCGCATCCGAGCGCGCTCGGGG	614		
QY	201	ProLeuValaGluGlnGlyArgValaArgAlaIaThrValGlySerLeuAlaGlyGlnPro	220		
DB	615	CCCCTGTGTGAACAGGGCGCGTGGCGCGCGCACTGTGGCTCCCTGGCGCGCAGCGG	674		
QY	221	LeuGlnGluArgAlaGlnAlaIaTrpGlyGluArgAlaArgMetGluGlnMetGly	240		
DB	675	CTACAGAGAGCGGGCGCGAGCGCTGGGGCGAGCGGTGCGCGGATGAGAGATGGGC	734		
QY	241	SerArgThrArgAspArgLeuAspGluValLysGluGlnValaIaGluValaArgAlaIys	260		
DB	735	AGCCGGACCCCGACCGCTGGAGCAGAGGTGAAGAGACAGGTGGAGGTGGCGCGCAGAG	794		
QY	261	LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaIaArgLeuLys	280		
DB	795	CTGGAGAGAGAGCCCGAGAGATACGCTGCGAGCGCGAGCGCTTCCAGGCCCGCGCTCAAG	854		
QY	281	SerTrpPheGluProLeuValaGluAspMetGlnArgGlnTrpAlaGlyLeuValaGluLys	300		

DB	855	AGCTGTTCCAGCCCTGTGTGGAGAGACATCAGCCGCTAGTGGCGCGGCTGTGTGAGAAAG	914		
QY	301	ValGlnAlaAlaValaGlyThrSerAlaIaIaProValaProSerAspAsnHis	317		
DB	915	GTGCAGCGTGGCGGTGGCGACAGCGCGCGCTGTGCTGCCAGGACAAATCAC	965		
RESULT 3					
E00823					
LOCUS					
DEFINITION	DNA sequence coding for human apolipoprotein E and its signal peptide.	1110 bp	DNA	linear	PAT 29-SEP-1997
ACCESSION	E00823				
VERSION	E00823.1	GI:2169084			
KEYWORDS	JP 1986096997-A/1.				
SOURCE	unidentified.				
ORGANISM	unidentified.				
REFERENCE	1 (bases 1 to 1110)				
AUTHORS	Teranishi,Y., Matsui,Y., Ikeda,Y. and Kimura,M.				
TITLE	PRODUCTION OF HUMAN APOLIPOPROTEIN E-LIKE PROTEIN				
JOURNAL	Patent: JP 1986096997-A 1 15-MAY-1986;				
COMMENT	MITSUBISHI CHEM IND LTD				
OS	Human (Homo sapiens)				
PN	JP 1986096997-A/1				
PD	15-MAY-1986				
PF	16-OCT-1984 JP 1984216987				
PI	TERANISHI YUTAKA, MATSUI YASUSHI, IKEDA YASUOKO, KIMURA MASAKO				
PC	CI2P21/00,A61K35/74,A61K37/04,CI2N15/00,(CI2P21/00,CI2R1:19),				
PC	(CI2N15/00,				
PC	CI2R1:19).				
CC	strandedness: Double;				
CC	topology: Linear;				
CC	hypothetical: No;				
CC	anti-sense: No;				
CC	*source: tissue_location: liver;				
FEH	key				
FEH	Location/Qualifiers				
FEH	3'UTR	1. 14			
FEH	sig_peptide	15. 68			
FEH	peptide'	product='human apolipoprotein E signal			
FEH	CDS	69. 968			
FEH	peptide'	product='human apolipoprotein E' FT			
FEH	FT	3'UTR			
FEATURES					
source	669. 1110.				
	Location/Qualifiers				
	1. 1110				
	/organism="unidentified"				
	/db_xref="taxon:32644"				
BASE COUNT	198 a	353 c	416 g	143 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	7.67e-84	Length:	1110		
Score:	1570.00	Matches:	315		
Percent Similarity:	99.37%	Conservative:	0		
Best Local Similarity:	99.37%	Mismatches:	2		
Query Match:	98.80%	Indels:	0		
DB:	6	Gaps:	0		
US-09-827-854-17 (1-317) x E00823 (1-1110)					
QY	1	MethylsValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal	20		
DB	15	ATGAAGGTTCTGTGGGCTGTGCTGTGCATTCCTGCGAGATGCCAGGCCAAGGTG	74		
QY	21	GIuGlnAlaValaGluThrGluProGluProGluLeuArgGlnGlnThrGlnTrpGlnSer	40		
DB	75	GAGCAACGGGTGGAGACAGAGCCGGAGCCGAGCTGCGCAGAGAGCCAGTGGCAGAC	134		
QY	41	GIyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr	60		
DB	135	GGCCAGCGCTGGGAACAGCGAGCTGGCTTTTGGATTACCTGCGGTGGCAGACA	194		

DB 855 ACCTGGTTCAGACCCCTGTGTGAAGACAGCAGCGCCATGGCGGCTGTGTGAGAG 914

QY 301 ValGlnAlaValaGlyThrSerAlaAlaProValProSerAspAsnHis 317

DB 915 GTGACAGCTGCCGTGGGACAGCGCGCCCTGTGTGCCAGACATTCAC 965

RESULT 3

LOCUS E00823

DEFINITION DNA sequence coding for human apolipoprotein E and its signal peptide.

ACCESSION E00823

VERSION E00823.1 GI:2169084

KEYWORDS JP 1986096997-A/1.

SOURCE unidentified.

ORGANISM unidentified.

REFERENCE 1 (bases 1 to 1110)

AUTHORS Teranishi, Y., Matsui, Y., Ikeda, Y. and Kimura, M.

TITLE PRODUCTION OF HUMAN APOLIPOPROTEIN E-LIKE PROTEIN

JOURNAL Patent: JP 1986096997-A 1 15-MAY-1986;

COMMENT MITSUBISHI CHEM IND LTD

OS Human (Homo sapiens)

PN JP 1986096997-A/1

PD 15-OCT-1984 JP 1984216987

PI TERANISHI YUTAKA, MATSUI YASUSHI, IKEDA YASUOKO, KIMURA MASAKO

PC C12P21/00, A61K35/74, A61K37/04, C12N15/00, (C12P21/00, C12N1:19),

PC (C12N15/00,

PC C12R1:19);

CC strandedness: Double;

CC topology: linear;

CC hypothetical: No;

CC anti-sense: No;

*source: tssuse-type=Liver;

key Location/Qualifiers

FT 3'UTR 1..14

FT sig-peptide 15..68

FT peptide' /product='human apolipoprotein E signal FT

FT CDS 69..968

FT /product='human apolipoprotein E' FT 3'UTR

FEATURES

source

location/Qualifiers

1..1110

/organism='unidentified'

/db_xref='taxon:32644'

BASE COUNT 198 a 353 c 416 g 143 t

ORIGIN

Alignment Scores:

Pred. No.: 7.67e-84

Score: 1570.00

Percent Similarity: 99.37%

Best Local Similarity: 99.37%

Query Match: 98.80%

Length: 1110

Matches: 315

Conservative: 0

Mismatches: 2

Indels: 0

Gaps: 0

US-09-827-854-17 (1-317) x E00823 (1-1110)

QY 1 MethylsValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20

DB 15 ATGAAGGTTCTGTGGGCTCGTGTGCTGCATTCCTGCGAGATGCCAGGCCAAGGTG 74

QY 21 GIuGlnAlaValaGluThrGluProGluProGluLeuArgGlnGlnThrGlnTrpGlnSer 40

DB 75 GAGCAAGCGGTGGAGACAGAGCCGGAGCTGCGCAGAGAGCCAGTGGCAGAGC 134

QY 41 GIyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60

DB 135 GGCCAGCCGCTGGGAAGTGCACCTGGCTTTTGGATTACCTGGCGGTGGCAGACA 194

	BASE COUNT	210 a	365 c	425 g	147 t	
ORIGIN						
Alignment Scores:						
Pred. No.:		7.94e-84			Length: 1147	
Score:		1570.00			Matches: 315	
Percent Similarity:		99.37%			Conservative: 0	
Best Local Similarity:		99.37%			Mismatches: 2	
Query Match:		98.80%			Indels: 0	
DB:		6			Gaps: 0	
US-09-827-854-17 (1-317) x AX302545 (1-1147)						
QY	1	MctLysValLeuTrpAlaAlaLeuIleuValThrPheLeuAlaGlyCysGlnAlaLysVal	20			
Dd	46	ATGAAGGTTCCTGTGGCGTGGCTGTGCATTCCTCGGAGGATGCCAGGCCAAGTTG	109			
QY	21	GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer	40			
Dd	106	GAGCAAGCGGTGGAGACAGACC CGGAGCCCAAGCTGCCCGCCAGCAGACCGAGTGCCAAGC	167			
QY	41	GlyGlnArgTrpGluLeuAlaLeuGlyValrPherPrAspTryLeuArgTrpValGlnThr	60			
Dd	166	GGCCAGCCCTGGAACTGGCACTGGGTGCTTTGGGATTACCTGGCCCTGGGTGACAACA	223			
QY	61	LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80			
Dd	226	CTGTCTGAGCAGGTGCCAGGAGAGACTGCTCATGTC CCAAGCTCACCCAGAACCTAGAGGGCG	285			
QY	81	LeuMetAspGluThrMetLysGluLeuLysAlaTryLysSerGluLeuGluGlnLeu	100			
Dd	286	CTGAATGGACGAGACCATGAAGAAGTTGAAGGCCATCAAATCTGCAACTGGAGAACAACTG	349			
QY	101	ThrProValAlaGluGluGlnThrArgAlaIarArgSerLysGluLeuGlnAlaAlaGlnAla	120			
Dd	346	ACCCGGGTGGGGAGAGAAACCGGGGACAGGCTGTTC AAGAGAGCTGCAGGGGGCCGACGGCC	409			
QY	121	ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTryArgGlyVal	140			
Dd	406	CGGCTGGGGCGGGACATGAGAGAGCTGTGGCGGGCCCTGGGTGCAGTACCGGGGGAGGTG	465			
QY	141	GlnAlaMetLeuAspGlnSerThrGluGluLeuArgValIArgLeuAlaSerHisLeuArg	160			
Dd	466	CAGGCCATGCTCTGGCCAAAGCACCGAGAGAGCTGGGGGTGGCCCTCCGCTCCACCTGCC	522			
QY	161	LysLeuArgLysArgGluLeuArgAspAlaAspAspleuGlnLysCysLeuAlaValTyr	180			
Dd	526	AAGCTGCTAAGCGGCTCTCCGCGATCCGATCACTGC GAGAGACCGCTGGCAGTGTAC	585			
QY	181	GlnAlaGlyAlaArgGluGlyValGlnIarGlyLeuSerAlaIleArgGluIarGluGly	200			
Dd	586	CAGGGCGGGGGCCCCGAGAGGGCGCGAGCGGGCTCTCA GCGCCATCCCAGCGCGCTGGGG	649			
QY	201	ProLeuValGluGlnGlyArgValIarGlnAlaIarThrValGlySerLeuAlaGlyGlnPro	220			
Dd	646	CCCCGTGGTAACAGGGCGGGCTCGGGCCGCCCACTGT GGGTCTCCCTGGCGCGCAACCG	709			
QY	221	LeuGlnGluIarGlnAlaIarTrpGlyGluIarGluLeuArgAlaIarGmetGluGluMetGly	240			
Dd	706	CTACAGGAGCGGGCCCAAGCGCTGGGGGAGCGCTGGCG GCGCGCGAGTAGAGAGATATGGC	765			
QY	241	SerArgTrpArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys	260			
Dd	766	AGCCGGACCCCGGACCGCCCTGGACGAGGTGTAAAGAC CAGGTGGCGGAGGTGCCGCCAAG	822			
QY	261	LeuGluGlnGlnAlaGlnGlnIleArgGluGlnAlaGluAlaPheGlnAlaArgLeuLys	280			
Dd	826	CTGGAGGAGCAGAGCCACCAATACATACGCTTCAGGCG CCGAGGCTTCACAGCCCGGCTCAAG	885			
QY	281	SerTrpPheGluIarProLeuValGluAspMetClnArgGlnTrpAlaGlyLeuValGlnLys	300			
Dd	886	AGCTGGTTCGAGCCCTGGTGGTAAGACATTCGACGCCA CAGTGGCGGGGCTGGGTGGAAG	949			

QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 946 GTGACAGCTCGCTGGGACACAGCGCCGCTGTGCTCCAGGACAAATAC 996

RESULT 5
 BD004278

LOCUS BD004278 1156 bp DNA linear PAT 31-JAN-2002
 DEFINITION Apo E humanized mammal
 ACCESSION BD004278
 VERSION BD004278.1 GI:18632239
 KEYWORDS JP 2001017028-A/2.
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 1 (bases 1 to 1156)
 Fujita, S., Hamanaka, H., Fukui, Y. and Yokoyama, M.
 Apo E humanized mammal
 Patent: JP 2001017028-A 2 23-JAN-2001;
 MITSUBISHI CHEMICAL CORP
 OS Homo sapiens (human)
 PN JP 2001017028-A/2
 PD 23-JAN-2001
 PF 28-APR-2000 JP 2000128919
 PR
 PI SHINOBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA
 A01K67/027, A61K45/00, A61P25/28, A61P43/00, C12N5/10, PC
 C12N15/09//C07K14/775,
 PC (C12N5/10, C12R1:91), C12N5/00, C12N15/00, (C12N5/00, C12R1:91) CC

FEATURES
 source FT CDS Location/Qualifiers
 1..1156
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 208 a 368 c 432 g 148 t
 ORIGIN

Alignment Scores:
 Pred. No.: 8.01e-84 Length: 1156
 Score: 1570.00 Matches: 315
 Percent Similarity: 99.37% Conservative: 0
 Best Local Similarity: 99.37% Mismatches: 2
 Query Match: 98.80% Indels: 0
 DB: 6 Gaps: 0

US-09-827-854-17 (1-317) x BD004278 (1-1156)

QY 1 MetLysValLeuThrAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 61 ATGAAGTCTGTGGGCTGCGTTCGTCGTCACATTCTTGGCAGATGCCAGGCGGAGTG 120

QY 21 GlnGlnAlaValGlnThrGlnProGlnLeuArgGlnGlnThrGlnThrGlnSer 40
 Db 121 GAGCAGACGGGTGGAGACAGCGGAGCCGAGCTCGCCAGAGAGACGAGTGCGGAGAC 180

QY 41 GlyGlnArgThrGlnLeuAlaLeuGlyArgPheThrAspTyrLeuArgTyrValGlnThr 60
 Db 181 GGCACAGCTGGGAACTGGCAGCTGGGTCTTTGGATTACCTGCGTGGGCGAGACA 240

QY 61 LeuSerGlnGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 Db 241 CTGTCTGACAGGTGACAGAGAGAGCTGCTCAGCTCCAGGTACCCAGAGACTGAGGGCG 300

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnGlnLeu 100
 Db 301 CTGATGAGAGACATGAAAGAGTTGAAGGCTTACAATCGAAGTGAAGAGACAACTG 360

QY 101 ThrProValAlaGlnGlnThrArgAlaAlaGluSerLysGlnGlnGlnAlaAlaGlnAla 120
 Db 361 ACCCCGCTGGCGAGAGACGCGGGCCAGCTGTCTCAAGAGACTGACGCGCGCGAGCC 420

QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
 Db 421 CGGCTGGGGCGGACATGAGAGACGTGTGGCGGCGCTGGTGCATGACCGGCGGAGTG 480

QY 141 GlnAlaMetLeuAspGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 481 CAGGCGCATGCTGGCGAGAGACAGGAGCTCGGGTGGCTCCCTCCCTCCACCTGCCG 540

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 Db 541 AAGCTCGTAAAGGCGCTCTCGCGGATGCCGATGACCTGCAGAAAGCCCTGCGAGTGC 600

QY 181 GlnAlaGlnAlaArgGlnGlnGlnAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
 Db 601 CAGCGCGGGCGCGGAGGCGCGGAGCGCGGCTTACGCGCCATCGCGGACGCGCTGGGG 660

QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 661 CCCCTGGTGGAAACAGGCGCGCTGCGGCGCGCCACTGTGGGCTCCCTGCGCGCGCAG 720

QY 221 LeuGlnGlnAlaArgAlaGlnAlaThrPglLysArgLeuArgAlaArgMetGlnGlnMetGly 240
 Db 721 CTACAGAGAGCGGCGCAGGCTGGGCGAGCGGCTGCGCGCGGATGGAGGATGGGC 780

QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
 Db 781 AGCCGACCGCGGACCGCGCTGGAGCGAGTGAAGAGACAGAGTGGCGGAGTGGCGGCAAG 840

QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 Db 841 CTGGAGAGACAGGCGCCAGAGATACCTGACGCGCGGAGCCCTCCAGCGCCCGCTCAAG 900

QY 281 SerThrPheGlnProLeuValGlnAspMetGlnArgGlnThrAlaGlyLeuValGlyLys 300
 Db 901 AGCTGGTTCGAGCCCTGTGTGAAGACATGACAGCGGCTGAGGCGGCTGTGTGGAGAA 960

QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 961 GTGACAGCTCGCTGGGACACAGCGCGCCCTGTGCTCCAGGACAAATAC 1011

RESULT 6
 HUMAPOE3 1156 bp mRNA linear PRI 24-NOV-2000
 LOCUS HUMAPOE3
 DEFINITION Homo sapiens preapolioprotein E (APOE) mRNA, complete cds.
 ACCESSION K00396
 VERSION K00396.1 GI:178850
 KEYWORDS apolioprotein; apolioprotein E; lipoprotein; polymorphism; very low density lipoprotein.
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 1 (bases 355 to 1156)
 Breslow, J.L., McPherson, J., Nussbaum, A.L., Williams, H.W., Lofquist-Kahl, F., Karathanasis, S.K. and Zannis, V.I.
 Identification and DNA sequence of a human apolioprotein E cDNA clone

JOURNAL J. Biol. Chem. 257 (24), 14639-14641 (1982)
 MEDLINE 83082756
 PUBMED 6897404

REFERENCE
 2 (bases 250 to 777)
 Wallis, S.C., Rogne, S., Gill, L., Markham, A., Edge, M., Woods, D., Williamson, R. and Humphries, S.
 The isolation of cDNA clones for human apolioprotein E and the detection of apoE RNA in hepatic and extra-hepatic tissues

JOURNAL EMBO J. 2 (12), 2369-2373 (1983)
 MEDLINE 84131952
 PUBMED 6199196

REFERENCE
 3 (bases 1 to 1156)
 Zannis, V.I., McPherson, J., Goldberger, G., Karathanasis, S.K. and Breslow, J.L.
 Synthesis, intracellular processing, and signal peptide of human apolioprotein E

281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300

Db 315 ACCCCGGTGGCGAGAGACGGCGGCACGGCTGTCCAAGAGAGCTGCAGCGCGCCAGGCC 374
 Oy 121 ATGLeuGlyAlaaspMetGluaspAlaCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 Db 375 CGGGTGGGGCGGACATGGAGAGACTGTGGCCCGCTGTGGTCTCACTACCGCGCGCAGAGTG 434
 Oy 141 GlnAlaMetLeuaspGlnSerThrGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 435 CAGGCGCATGTCTGGCCACAGCAGCAGCGAGCTGGCGGCTGGCTCCACTGGCGC 494
 Oy 161 LysLeuArgGlyArgLeuLeuArgaspAlaaspAspLeuGlnLysCysLeuAlaValTyr 180
 Db 495 AAGCTGCTTAAGCGGCTCTCCGCGATGCCATGTGCATGCAGAAAGCGCTGGCACTGTAC 554
 Oy 181 GlnAlaGlyAlaArgGlyGluValGlnArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db 555 CAGGCGGGGGCGCGGAGGGCGCGAGCGCGCTGCACGCGCCATCCGCGAGCGGCTGGGG 614
 Oy 201 ProLeuValGlnGlnGlyArgValArgValAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 615 CCCCTGGTGGAAACAGCGCGCGGTGGCGCGCACTGTGGCTCCCTGGCGCGCCAGCGC 674
 Oy 221 LeuGlnGluArgAlaGlnAlaTyrGlyLeuArgLeuArgAlaArgMetGlnGluMetGly 240
 Db 675 CTACAGGAGCGCGCGCCAGCGCTGGGGCGAGCGCGCTGGCGCGATGAGAGAGATGGGC 734
 Oy 241 SerArgThrArgaspArgLeuaspGlnValLysGlnGlnValAlaGluValArgAlaLys 260
 Db 735 AGCCCGAGCCCGGACCGCGCTGGAGAGGTGAAGAGCAGGTGGCGGAGGTGGCGCCCAAG 794
 Oy 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 Db 795 CTGGAGGAGCAGCGCCACAGATACGCTGCAGCGCGAGCGCTTCACAGCCCGCTCAAG 854
 Oy 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
 Db 855 AGCTGTGTCGAGCCCTGCTGTGAGACATGCAGCGCCAGTGGCGCGCTGTGGAGAG 914
 Oy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 915 GTGCAGGCTGCCGTGGGACCAAGCGCGCCCTGTGTGCCAGGACAAATCAC 965
 RESULT 9
 AX333278 1157 bp DNA linear PAT 09-JAN-2002
 LOCUS AX333278
 DEFINITION Sequence 3787 from Patent W00194629.
 ACCESSION AX333278
 VERSION AX333278.1 GI:18123912
 KEYWORDS
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 1
 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
 Horrikan, S., Soppet, D.R., and Weaver, Z.
 Cancer gene determination and therapeutic screening using signature
 gene sets
 JOURNAL Patent: WO 0194629-A 3787 13-DEC-2001;
 AVALON Pharmaceuticals (US)
 FEATURES
 source
 1. .1157
 Location/Qualifiers
 BASE COUNT 212 a 370 c 426 g 149 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.69e-83 Length: 1157
 Score: 1561.00 Matches: 313
 Percent Similarity: 98.74% Conservative: 0
 Best Local Similarity: 98.74% Mismatches: 4
 Query Match: 98.24% Indels: 0

DB: 6 Gaps: 0
 US-09-827-854-17 (1-317) x AX333278 (1-1157)
 Oy 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 62 ATGAAGGTTCTGTGGGCTGGCTGGTGTGCATCTCTGGCAGAGATGGCCAGCCAAAGGTG 121
 Oy 21 GlnGlnAlaValGlnTrpGluProGluProGluLeuArgGlnGlnTrpGlnTrpGlnSer 40
 Db 122 GAGCAAGCGGTGGAGACAGACCGGAGCCGAGCTGCCACAGCAGACCGAGTGGCAGAGC 181
 Oy 41 GlyGlnArgTrpGlnLeuAlaLeuGlnLysArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 Db 182 GGCCAGCCCTGGGAACCTGGCACCTGGGTGCTTTTGGGATTACCTGGCTGGGTGGAGACA 241
 Oy 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 Db 242 CTGTCTGAGCAGGTGCAGAGAGAGCTGCTCAGCTCCCAAGTCACCCAGAACTAGAGGCGC 301
 Oy 81 LeuMetAspGlnTrpMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnGlnLeu 100
 Db 302 CTGATGACGAGACCATGAGAGAGTTGAAGGCTTCAAAATGTGGAACCTGGAGAACACTG 361
 Oy 101 ThrProValAlaGlnGlnGlnTrpArgAlaArgLeuSerLysGlnGlnAlaAlaGlnAla 120
 Db 362 ACCCGGTAGCGAGAGAGAACCGCGGACGCTGTCCAAGAGACTGCACAGCGCGAGGCC 421
 Oy 121 ArgLeuGlnAlaaspMetGlnaspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 Db 422 CGCGTGGCGCGGACATGAGAGAGCTGTGGCGCGCTGTGGTGTGAGTACCGCGCGAGGTG 481
 Oy 141 GlnAlaMetLeuaspGlnSerThrGlnGlnGlnValArgValArgLeuAlaSerHisLeuArg 160
 Db 482 CAGGCGATGCTCGCGCAGAGCAGCAGAGAGCTGGCGGTGGCTGTCCCTCCACTGTGGC 541
 Oy 161 LysLeuArgGlyArgLeuLeuArgaspAlaaspAspLeuGlnLysCysLeuAlaValTyr 180
 Db 542 AAGCTGCTTAAGCGGCTCTCCGCGATGCCATGTGCAGAGAGCGCTTGGCACTGTAC 601
 Oy 181 GlnAlaGlyAlaArgGlnGlnGlnValArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db 602 CAGCGCGGGCGCGGAGGGCGCGAGCGCGGCTGCACGCGCATCCGCGAGCGCTGGGG 661
 Oy 201 ProLeuValGlnGlnGlnGlyArgValArgValAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 662 CCCCTGTGGAAACAGCGCGGTGGCGCGCCACTGTGTGGCTCTCCGTGGCGCGCCAGCGC 721
 Oy 221 LeuGlnGlnArgAlaGlnAlaTyrGlnGlnArgLeuArgAlaArgMetGlnGluMetGly 240
 Db 722 CTACAGGAGCGGGCCACAGCCTGGGGCGAGCGGCTGGCGCGGATGGAGAGATGGGC 781
 Oy 241 SerArgThrArgaspArgLeuaspGlnValLysGlnGlnValAlaGluValArgAlaLys 260
 Db 782 AGTCGAGACCCCGACCGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 841
 Oy 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 Db 842 CTGGAGGAGCAGCGCCACAGATACGCTTGCAGCGCGCTTCCAGCGCGCTCTCAAG 901
 Oy 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
 Db 902 AGCTGTGTCGAGCCCTGCTGTGAGACATGTGCAGCGCCAGTGGCGCGCTGTGGAGAGAG 961
 Oy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 962 GTGCAGGCTGCCGTGGGACCAAGCGCGCCCTGTGTGCCAGGACAAATCAC 1012
 RESULT 10
 AX409597 1157 bp DNA linear PAT 14-JUN-2002
 LOCUS AX409597
 DEFINITION Sequence 2244 from Patent W00229103.
 ACCESSION AX409597

VERSION	AA409597.1	GI:21442302
KEYWORDS	human.	
SOURCE	ORGANISM Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1	
AUTHORS	Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.	
TITLE	Gene expression profiles in liver cancer	
JOURNAL	Patent: WO 0229103-A 2244 11-APR-2002;	
FEATURES	GENE LOGIC INC (US)	
SOURCE	Location/Qualifiers	
	1..1157	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/note="EMBL/Genbank Accession No. M12529"	
BASE COUNT	212 a	370 c 426 g 149 t
ORIGIN		
Alignment Scores:		
Pred. NO.:	2.69e-83	Length: 1157
Score:	1561.00	Matches: 313
Percent Similarity:	98.74%	Conservative: 0
Best Local Similarity:	98.74%	Mismatches: 4
Query Match:	98.24%	Indels: 0
DB:	6	Gaps: 0
US-09-827-854-17 (1-317) x AX409597 (1-1157)		
QY	1	MeltySVAlleuTTPAlaAlaLeuValThrPheLeuAlaGlycysGlnAlaLysVal 20
DB	62	ATGAAGGTCTGTGGGCTGCTGTGTCATCTCGCAGATGCCAGGCCAAGTG 121
QY	21	GlUGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTPGlnSer 40
DB	122	GAGCAAGCGGTGGAGACAGCCCGAAGCCAGCTGCGCCAGCAGACCGAGTGGCAGAGC 181
QY	41	GlyGlnArgTPGluLeuAlaLeuGlyArgPheThrPaspTryLeuArgTPValGlnThr 60
DB	182	GGCCAGCGCTGGGAACGTGCACCTGGGTCTTTGGGATTACCTGGCGTGGTGCAACA 241
QY	61	LeuSerGluGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB	242	CTGTCTGAGCAGGTGCGACAGAGACTGCTGCATGCCAAGTACCACGAAGATGAGGGCG 301
QY	81	LeuMetAspGluThrMetLysGluLeuLysAlaTryLysSerGluLeuGlnGluGlnLeu 100
DB	302	CTGATGGACGACCATCGATGAGAGATTGAAGGCTTCAACAATCGAAGTGGAGAACACTG 361
QY	101	ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB	362	ACCCCGGTAGCGGAGAGACCGGGCACGGGTGTGTCAAAGGAGGTGCAGAGCGCGCAGGCC 421
QY	121	ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTryArgGlyGluVal 140
DB	422	CGGCTGGGCGCGGACATGAGAGAGCTGTGGCGCCCTGTGGTGCAGTACCGGGGAGGTG 481
QY	141	GlnAlaMetLeuAspGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB	482	CAGGCCAAGTCTGGCCAGACAGCACCGAAGAGCTGGGGGTGGCCCTGCCACACTTCGC 541
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
DB	542	AAGTCGCTAAGCGGCTCTCTCCGGATCCGATATACCTGCCAAGAGCGCTGGCAGTTAC 601
QY	181	GlnAlaGlyAlaArgGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB	602	CAGCGCGGGGCCCGCGAGGGCGCGGAGCGGCTTCAGAGCGCATCCGCGAGCGCTGGGG 661
QY	201	ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB	662	CCCTGTGTGGTAACAGGGCGCGTGTGGGGCCGCCACTGTGGGTCTCCCTGGCGGGCACCG 721

Oy	221	LeungLingUatAgaLaGInAlaATrPcJylUarGleuArGaLaATgMeGlUmecSly	240
Db	722	CTACAGGACCGGGCCCAAGCCTGGGGCGAGGGCTGCCGGCGGAGATGAGGAATGGGC	781
Oy	241	SerTrpArgTrtArgAspArgLeuAspPolValIylSGluGlnValAlaGlnValArgAlas	260
Db	782	AGTCGGACCCTGGACCGCTGTGGACGAGTGAAGACAGAGTGCGGAGGTGCGGCCAAG	841
Oy	261	LeugLingUlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPhagInAlaArgLeuls	280
Db	842	CTGGAGGAGCAGAGCCCAACAGATACCGCTGCAGAGGCCAGGCTTCAGAGCCGCTCAAG	901
Oy	281	SerTrpPheGluProLeuValGIuAspMetLnaArgIntPrAlaGlyLeuValGIuIs	300
Db	902	AGCTGGTTCCGAGCCCTGGTGGAAAGCATGCAAGCGCCAGTGGCGCGGCTGTGGAGAAG	961
Oy	301	ValGlnAlaAlaValGIYThrSerAlaIalProValProSerAspAsnHis	317
Db	962	GTCGAGGCTGCCTGGTGGCACAGCGCCGCTGTGCTTGCACGACAAATTCAC	1012
RESULT 11			
LOCUS	115975	115975	1157 bp DNA linear PAT 03-APR-1996
DEFINITION	Sequence 1 from patent US 5472858.		
ACCESSION	115975		
VERSION	115975.1	GI:1250883	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1157)		
AUTHORS	Attie,A.D., Gretch,D.G., Sturley,S.L. and Beckage,N.E.		
TITLE	Production of recombinant proteins in insect larvae		
JOURNML	Patent: US 5472858-A 1 05-DEC-1995;		
FEATURES	Location/Qualifiers		
source	1..1157		
BASE COUNT	212 a	370 c	426 g 149 t
ORIGIN			
Alignment Scores:			
Pred. NO.:	2.69e-83	Length:	1157
Score:	1561.00	Matches:	313
Percent Similarity:	98.74%	Conservative:	0
Best local Similarity:	98.74%	Mismatches:	4
Query Match:	98.24%	Indels:	0
Db:	6	Gaps:	0
US-09-827-854-17 (1-317) x 115975 (1-1157)			
Oy	1	MettySalLeutTPPAAlaLeuLeuValThrpHeLeuAlaGIycSGlnAlaIysVal	20
Db	62	ATGAAGATTCTGTGGGCTGCTGTGTCAATTCTCGGAGGATCCAGGCCAAGGTG	121
Oy	21	GIuGlnAlaValGIuThrgLuProGluProGIuLeuArgGlnIthrGIUTripGlnSer	40
Db	122	GAGCAAGCGGTGGAGACAGACCAGGACCCAGCTGCCCGCCAGCAGACCGAATGTGCAGAGC	181
Oy	41	GIYGLnArgrTPGIuLeuAlaLeuGIYArphetrPrpsAprrYLearArgrTPvaIGlnThr	60
Db	182	GGCCAGCCCTGGGAACCTGCACCTGGGTGCTTTGGGATTAACCTGCCTGGGTGCAGACA	241
Oy	61	LeuSerGIuGlnValGIuGlnGIuLeuLeuSerSerGlnValIThrGlnIuLeuArgAla	80
Db	242	CTGTCTGACGACGGTGCAGAGAGACTGTCTACAGCTCCCAAGTCACCCCAAGAACCTAGGGCG	301
Oy	81	LeuMetAspGIuThrMetIysGIuLeuIysAlaTYrlYrsSerGIuLeuGIuGlnIuLeu	100
Db	302	CTGATGACGACGACCATGAAGAGATTGAAGCCTTAACAATCGGAACCTGAGAACCACTG	361
Oy	101	ThrpProValAlaGIuThrArGIAlaArgLeuSerIysGIuLeuGlnAlaAlaGlnAla	120
Db	362	ACCCCGTAGCGGAGAGACCGGGCAACGGCTGTCCAAAGAGAGCTGCACAGACGCGCCAGGCC	421

QY 121 ATGLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 DB 422 CGGCGGGGGGGGACATGAGAGACGTGTCCGGCCCGCTGTCTACATACCGCGGAGG 481
 QY 141 GlnAlaMetLeuAspGlnSerThrGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 482 CAGGCGCATCTCTGGCCAGCACCAGGAGCTGGCGGTCCGCTCCACCTCCGCGC 541
 QY 161 LysLeuArgValArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 DB 542 AAGCTGCGGAGCGGCTCTCCGCGATCCCGATGACCTGCAGAACCGCTGTGCACTGAC 601
 QY 181 GlnAlaGlyAlaArgGluGlyValArgGlyLeuSerAlaIleArgGlyLeuArgGly 200
 DB 602 CAGGCGGGGGGGCGGAGGCGCCGAGCGCGCTCAGCGCCATCCGCGAGCGCTGGG 661
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 DB 662 CCCCTGTGTGAGACAGGCGCGCTGGCGCGCCACTGTGTGCTCCCTGGCGCGCAGCG 721
 QY 221 LeuGlnGluArgAlaGlnAlaThrPglGluArgLeuArgAlaArgMetGluGlnMetGly 240
 DB 722 CTACAGGAGCGGGCCAGGCGCTGGGGCGAGCGCTGGCGCGGATGAGAGATGGCG 781
 QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGluValArgAlaLys 260
 DB 782 ACTCGGAGCCCGGACCGCTGTGACAGAGTGAAGAGACAGTGGCGGCGCCGCAAG 841
 QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 DB 842 CTGGAGAGAGCAGCCAGCAGATACGCCCTGAGCGCGAGCGCTCCAGGCGCGCTCAAG 901
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 DB 902 AACTGTTGTGACCCCTGTGTGAGACATGACGCCCATGTCGGGGCGGTGTGTGAGAG 961
 QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 962 GTGACAGGCGCGCTGGGACACAGCGCGCGCTGTGTGCCAGGACATATAC 1012
 RESULT 12
 HUMAPOE
 LOCUS HUMAPOE 1157 bp mRNA linear PRI 08-AUG-1995
 DEFINITION Human apolipoprotein E mRNA, complete cds.
 ACCESSION M12529
 VERSION M12529.1 GI:178848
 KEYWORDS apolipoprotein.
 SOURCE Homo sapiens (clone: PHAE[112,178,813].) male 57-year old liver
 CDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1157)
 AUTHORS McLean,J.W., Elshourbagy,N.A., Chang,D.J., Mahley,R.W. and
 Taylor,J.M.
 TITLE Human apolipoprotein E mRNA. CDNA cloning and nucleotide sequencing
 of a new variant
 JOURNAL J. Biol. Chem. 259 (10), 6498-6504 (1984)
 MEDLINE 84212473
 PUBMED 6337682
 FEATURES
 source
 1..1157 location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="19q13.2"
 /clone="PHAE[112,178,813]."
 /sex="male"
 /tissue_type="liver"
 /dev_stage="57-year old"
 1..1157
 /gene="APOE"
 mRNA

CDS
 /gene="APOE"
 /note="apoe mRNA; G00-119-691"
 62..1015
 /gene="APOE"
 /note="precursor"
 /codon_start=1
 /product="apolipoprotein E"
 /protein_id="AAB59518.1"
 /db_xref="GI:178849"
 /db_xref="GDB:G00-119-691"
 /translation="MKVIMALLVTFLLAGCAKVEQAVETPEPEPLRQTEMOSQRM
 ELAIGREWDYLRWQVTSQVQVEELSSQVQELRALMDETREKELKAVSLERDLP
 VAETFRARLSEKLEOTQARILQADMEWCGRLVOYRGVQAMIGOSTELRYRLAHLR
 KLRRLDLPDLOKRLAVYQAGAREGERLSAIRELIGPLVEQGRVRAATVSLAG
 OPLERAAQMEERLIRARMEEGSRTRDLDVKKRQVAREVRARLEBQAQIRLQAEAFQ
 ARLSMPEPLVEDMDQRMQAGLVEKVAVGTSAPVPSDNH"
 62..115
 sig_peptide
 /gene="APOE"
 /note="G00-119-691"
 116..1012
 mat_peptide
 /gene="APOE"
 /product="apolipoprotein E"
 /note="G00-119-691"
 370
 variation
 /gene="APOE"
 /note="a in PHAE[112,178]; g in PHAE813; G00-119-691"
 /replace="g"
 410
 variation
 /gene="APOE"
 /note="a in PHAE[112,178]; g in PHAE813; G00-119-691"
 /replace="g"
 569
 variation
 /gene="APOE"
 /note="c in PHAE[112,178]; g in PHAE813; G00-119-691"
 /replace="g"
 784
 variation
 /gene="APOE"
 /note="t in PHAE[112,178]; c in PHAE813; G00-119-691"
 /replace="c"
 859
 variation
 /gene="APOE"
 /note="g in PHAE[112,178]; a in PHAE813; G00-119-691"
 /replace="a"
 BASE COUNT 212 a 370 c 426 g 149 t
 ORIGIN 30 bp upstream of BstNI site.
 Alignment Scores:
 Pred. No.: 2.69e-83 Length: 1157
 Score: 1561.00 Matches: 313
 Percent Similarity: 98.74% Conservative: 0
 Best Local Similarity: 98.74% Mismatches: 4
 Query Match: 98.24% Indels: 0
 DB: 9 Gaps: 0
 US-09-827-854-17 (1-317) x HUMAPOE (1-1157)
 QY 1 MetLysValLeuThrAlaAlaLeuValTrpPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 62 ATGAAGTTCTGTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 121
 QY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGlnTrpGlnSer 40
 DB 122 GAGCAAGCGGTGTGAGACAGCAGCGGAGCGCGGAGCTGTGCGCACAGACCGAGTGGCAGAC 181
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB 182 GGCACAGCGCTGTGAGACTGTGACATGTGGTCTTTTGGATCTGCGTGGGTGTGACAGCA 241
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB 242 CTGTGTGACAGCTGTGAGAGGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 301
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100

|||||
Db 302 CTGATGACAGACACATGAAAGAGTTGAMGGCTTACAAATCGAAGCTGGGAGACACTG 361
QY 101 TheProVa1A1aG1uG1uThrArGa1AaRgLeuSerLyS1uG1uG1n1a1aG1n1a1a 120
|||||
Db 362 ACCCGGTTAGCGAGAGAGACGGGGCAGCGCTCTCCAGAGAGTGCAGACGGCGAGGCC 421
QY 121 ArgLeuG1yAlaAspMetG1uAspValCysG1yArgLeuValG1nTyrArgG1yG1uVal 140
|||||
Db 422 CGGCTGGGCGCGGACATGAGGACGTGTGGCGCGCTGTGTCAGTACCCCGCGAGAGTG 481
QY 141 G1nAlaMetLeuAspG1nSerThrG1uG1uLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 482 CAGGCGATGCTCGGCGAGACACCGAGAGACTCGGGTGGCTCGCTCCCTCCACCTGGCG 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuG1nTyrCysLysAlaVal1Tyr 180
|||||
Db 542 AAGCTGCGTAAGCGGCTCTCCCGCATCCGATGACTGCAGAAAGCGCTGGCAGTGTAC 601
QY 181 G1nAlaG1yAlaAaRg1uG1uG1yAlaAaRg1yLeuSerAla1LeaRg1uAaRgLeuG1y 200
|||||
Db 602 CAGGCGGGGCGCGGAGGCGCGGAGCGGCTCAGCGCATCCGCGAGCGCTGGGG 661
QY 201 ProLeuValG1uG1nG1yArgValArgAlaAlaThrValG1ySerLeuAlaG1yG1nPro 220
|||||
Db 662 CCCCTGTGGAACAGGAGCGCGGTGGGCGCCACGTGTGGCTCTCGCGCGGCGAGCGG 721
QY 221 LeuG1nG1uAaRg1uAaRg1uAaRg1yG1yG1uAaRg1uAaRgMetG1uG1uMetG1y 240
|||||
Db 722 CTACAGAGAGGGGCGCCAGGCGCTGGGGCGAGCGGCTGGCGCGGATGGAGAGATGGGG 781
QY 241 SerArgThrArgAspArgLeuAspG1uValLysG1uG1nValAlaG1uValArgAlaLys 260
Db 782 AGTCGAGACCGCGCACCGCTCGGACGAGGTGAAGACAGAGTGGCGGAGGTGGCGCGCAAG 841
QY 261 LeuG1uG1uG1nG1uAaRg1uG1n1LeaRgLeuG1nAlaG1uAaRg1uAaRg1uAaRgLeuLys 280
|||||
Db 842 CTGGAGAGAGAGCGCCACGACGATACCGCTGCAGGCGGAGGCTTCCAGGCGCGCTCAAG 901
QY 281 SerTrpPheG1uProLeuValG1uAspMetG1nAaRgG1nTrpAlaG1yLeuValG1uLys 300
|||||
Db 902 AGCTGTGTCAGGCGCCCTGGTGGAAACATGACAGCGGCAGTGGCGGGGTGGTGGAGAA 961
QY 301 ValG1nAlaAlaValG1yThrSerAlaAlaProValProSerAspAsnHis 317
|||||
Db 962 GTGCAGAGCTGCCTGGCGCACACGCGCGCTGTGCTCCAGCGACATGAC 1012
RESULT 13
BD004279 1156 bp DNA linear PAT 31-JAN-2002
LOCUS BD004279 Apo E humanized mammal.
DEFINITION
ACCESSION BD004279.1 GI:18632240
VERSION BD004279.1 GI:18632240
KEYWORDS JP 2001017028-A/3.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1156)
Fujita, S., Hamanaka, H., Fukui, Y. and Yokoyama, M.
Apo E humanized mammal
Patent: JP 2001017028-A 3 23-JAN-2001.
MITSUBISHI CHEMICAL CORP
OS Homo sapiens (human)
PN JP 2001017028-A/3
PD 23-JAN-2001
PE 28-APR-2000 JP 2000128919
PR SHINOBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA PC
A01K67/027, A61K45/00, A61P25/28, A61P43/00, C12N5/10, PC
C12N15/09, C07K14/775,
PC (C12N5/10, C12R1:91), C12N5/00, C12N15/00, (C12N5/00, C12R1:91) CC

Alignment Scores:
Pred. No.: 4,03e-83 Length: 1156
Score: 1558.00 Matches: 314
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 99.05% Mismatches: 3
Query Match: 98.05% Indels: 0
Gaps: 0
US-09-827-854-17 (1-317) x BD004279 (1-1156)
QY 1 MetLysValLeuTrpAlaLeuLeuVal1ThrPheLeuAlaG1yCysG1nAlaLysVal 20
|||||
Db 61 ATGAAGTTCTGTGGGCTCGTTGCTGTGCATTCCTCGGAGATGCCAGGCAAGGTG 120
QY 21 G1uG1nAlaValG1uThrG1uProG1uProG1uLeuArgG1nThrG1uTrpG1nSer 40
|||||
Db 121 GAGCAAGCGGTGAGACAGAGCGGAGCGCGAGCTGCGCACAGACAGAGTGGAGAC 180
QY 41 G1yG1nArgTrpG1uLeuAlaLeuG1yArgPheTrpAspTyrLeuArgTrpValG1nThr 60
|||||
Db 181 GGCACAGCTGGGAACCTGGACACTGGGTGCTTTGGGATTACTCGCTGGGTGGAGACA 240
QY 61 LeuSerG1uG1nValG1nG1uLeuLeuSerSerG1nVal1ThrG1uLeuArgAla 80
Db 241 CTGTCTGACAGAGTGCAGAGAGAGCTGCACCTCCAGTCCAGTCCAGGAACTGAGGGG 300
QY 81 LeuMetAspG1uThrMetLysG1uLeuLysAlaTyrLysSerG1uLeuG1uG1nLeu 100
Db 301 CTGATGACAGACCACTGAAGAGTTGAAGGCTTACAAATCGGAAGTGAAGACACTG 360
QY 101 TheProVa1A1aG1uG1uThrArGa1AaRgLeuSerLyS1uG1uG1n1a1aG1n1a1a 120
|||||
Db 361 ACCCGGTTAGCGAGAGAGACCGGGCAGCGGTGTCACAGAGCTGCAGGCGCGAGGCC 420
QY 121 ArgLeuG1yAlaAspMetG1uAspValCysG1yArgLeuValG1nTyrArgG1yG1uVal 140
|||||
Db 421 CGGCTGGGCGCGGACATGAGAGACGTGCCGCGCTGTGCTCAGTACCGCGCGAGGTG 480
QY 141 G1nAlaMetLeuAspG1nSerThrG1uG1uLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 481 CAGGCGATGCTGGCGCACAGACACCGAGAGCTGCGGGTCCGCTCCACCTCGGCC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuG1nLysCysLeuAlaVal1Tyr 180
|||||
Db 541 AAGCTGCTTAAGCGGCTCTCCGCGATGCCATGACCTGCAGAAAGCGCTGCAGTGTAC 600
QY 181 G1nAlaG1yAlaAaRg1uG1uG1yAlaAaRg1yLeuSerAla1LeaRg1uAaRgLeuG1y 200
|||||
Db 601 CAGGCGGGGCGCGGAGGCGCGGAGCGGCTCAGGCGCATCCGCGAGGCGCTGGGG 660
QY 201 ProLeuValG1uG1nG1yArgValArgAlaAlaTrpValG1ySerLeuAlaG1yG1nPro 220
|||||
Db 661 CCCCTGTGGAACAGGCGCGCTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 221 LeuG1nG1uAaRg1uAaRg1uAaRg1yG1yG1uAaRg1uAaRgMetG1uG1uMetG1y 240
|||||
Db 721 CTACAGAGAGCGGCGCCAGGCTGGGCGAGCGGCTGCGCGCGGAGTGGAGTGGGC 780
QY 241 SerArgThrArgAspArgLeuAspG1uValLysG1uG1nValAlaG1uValArgAlaLys 260
|||||
Db 781 ACCCGGACCGGAGCGCTGTGAAGACAGGTGAAGAGCACTGGAGGTGGCGCGCAAG 840
QY 261 LeuG1uG1uG1nG1uAaRg1uG1n1LeaRgLeuG1nAlaG1uAaRg1uAaRgLeuLys 280
|||||

Db 841 CTGGAGGACGAGCCAGCATACGCCCTGCAGGCCGCTCCAGGCCGCTCCAG 900
QY 281 SerTpphGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluIys 300
Db 901 ACTGTTGACCCCTGGTGAAGCATGCAGCCGCACTGGGCGGCTGGTGGAGAG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCGGTGGGACGAGCGGCCCTGTGGCCAGGACATCATC 1011
RESULT 14
MEAPOE 1178 bp mRNA linear PRI 31-MAR-1995
LOCUS Monkey mRNA for apolipoprotein E.
DEFINITION X13887
ACCESSION X13887
VERSION X13887.1 GI:38054
KEYWORDS apolipoprotein; apolipoprotein E.
SOURCE Macaca fascicularis.
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopitheciinae; Macaca.
REFERENCE 1 (bases 1 to 1178)
AUTHORS Marotti K.R.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1988) Marotti K.R., The Upjohn Co, 7242 209 713,
301 Henrietta Street, Kalamazoo, MI 49008
2 (bases 1 to 1178)
AUTHORS Marotti K.R., Whitted, B.E., Castle, C.K., Polites, H.G. and
Weichlor, G.W.
TITLE Nucleotide sequence of the cynomolgus monkey apolipoprotein E cDNA
JOURNAL Nucleic Acids Res. 17 (4), 1778 (1989)
MEDLINE 89160349
PubMed 2922300
COMMENT Data kindly reviewed (20-Mar-1989) by Marotti K.R.
FEATURES
source Location/Qualifiers
1..1178
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="1"
/tissue_type="liver"
/clone_1lb="lambda gt11"
83..1036
/codon_start=1
/product="apolipoprotein E (AA -18 to 299)"
/protein_1lb="CAA32092.1"
/db_xref="GI:38054"
/db_xref="SWISS-PROT:P10517"
/translation="MKVLMALVTFILGCOAKYQEPVETPELROAGSGSGPM
ELALGRFMDYLRWOTLSQVDELLSPQVOTLTMDETMKELKAKSKLEFOLSP
VAEETRARLSKELQAAOARIGADMEDVSRIVORYSEVQAMLGSTEBELRLASHLR
KLKRLRLDADDLQRLAVYQAGAREGEVSAIREFLGLVEGGRVRAIVGSLAS
QPLDERAQLGEIRLARKMEGSSRTRDLDVEKQVAEVRALKEQAOQISLQAEAFQ
ARLKSEPFLEVDHQRQWAGLVEKVOAAVGAATAVPIDNH"
83..136
sig_peptide 137..1033
mat_peptide /product="apolipoprotein E (AA 1 - 299)"
1154..1159
misc_feature /note="polyA signal"
BASE COUNT 217 a 373 c 426 g 162 t
ORIGIN

Alignment Scores:
Pred. No.: 5.67e-76 Length: 1178
Score: 1436.00 Matches: 294
Percent: 93.38% Conservative: 2
Best Local Similarity: 92.74% Mismatches: 21
Query Match: 90.37% Indels: 0
DB: 9 Gaps: 0

US-09-827-854-17 (1-317) x MEAPOE (1-1178)
QY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20

Db 83 ATGAAGGTTCTGGGCTCGTTGCTGTCACATTCCTGGCAGATGCCAGGCAAGTG 142
QY 21 GluGlnAlaValGlyThrLeuProGluProGluLeuArgGlnGlnThrLeuTrpGlnSer 40
Db 143 GAGCAACCGGTGGAGCCAGCAAGAGAGGAAACCGAGCTTCGACACAGGCTGAGGCCAGAC 202
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
Db 203 GGCACGCCCTGGGAGCTGGACATGGGTCCCTTTTGGATTAATCTGCTGGGTGCAGACA 262
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnLeuLeuArgAla 80
Db 263 CTGTCTGACACAGTCGACAGAGAGAGCTGCTCACGCCCCAGCTCACACAGAACTGACAGCAG 322
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGluGlnLeu 100
Db 323 CTGATGGAGACCAATGATGAGAGATTGAAGGCCCTCAAAATCGAACTGAGAGAACATG 382
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 383 ACCCGGTGGCGAGAGAGACGCGGCGCATGCTCTCAAGAGACTGCAAGCGGCGCAGGCC 442
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
Db 443 CGGTGGGTGGTCCGACATGAGAGACGTCGACAGCGGCTGTGTCAGATGACCGACAGCGAGTG 502
QY 141 GlnAlaMetLeuAspGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 503 CAGGCCATCTGGGCGACGATGACCGAGCGAGCGGCGGCGGCTGCTGCCATCCACCTGCGCC 562
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 563 AAGCTGCGCAACCGGCTCTCCGCGATGCTGATGACCTGCACAAAGCGCTGCGCAGTGTT 622
QY 181 GlnAlaGlyAlaArgGluGluGluValGluArgGlyLeuSerAlaTrpArgGluArgLeuGly 200
Db 622 CAGCGCGGGGCGCGGAGAGCGCGCGCGGCGGCGGATCGCATCCGACAGCGCGCTCGGGA 682
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyLysPro 220
Db 683 CCCCTGTGTGACAGAGCGCGCTGCGGCGGCGGCGGCTGCTGCTGCGGCGGCGGCGGCGG 742
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluLysGly 240
Db 743 CTTCAAGAGCGGCGCGAGGCTTGGGTGAGCGGCTTGGCGCGGATGAGAGATGGGCG 802
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 803 ACCCGGACCGCGGACCGCTGACGAGTGAGAGAGCAGGTGGCGAGGTGGCGCGCAAG 862
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 863 CTGGAGGAACAGCGCCAGACAGATAGCCCTGACGCCGAGCGGCTTCACAGCCCGCTCAAG 922
QY 281 SerTpphGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluIys 300
Db 923 ACTGTTGACCCCTGGTGAAGATATGACGCCCACTGGGCGGCTGGTGGAGAG 982
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 983 GTGCAGGCTGCGGTGGGCGGCGGCGGCCCTGTGGCCAGGACATCATC 1033
RESULT 15
AF261279 5491 bp DNA linear PRI 27-OCT-2000
LOCUS Homo sapiens apolipoprotein-E gene, complete cds.
DEFINITION AF261279
ACCESSION AF261279
VERSION AF261279.1 GI:11034800
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 5491)
AUTHORS Nickerson,D.A., Taylor,S.L., Fullerton,S.M., Weiss,K.M.,
Clark,A.G., Stengard,J.H., Salomaa,V., Boerwinkle,E. and Sing,C.F.
TITLE Sequence diversity and large-scale typing of SNPs in the human
apolipoprotein E gene
JOURNAL Genome Res. 10 (10), 1532-1545 (2000)
MEDLINE 2049366
PubMed 11042151
REFERENCE 2 (bases 1 to 5491)
AUTHORS Nickerson,D.A.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2000) Department of Molecular Biotechnology,
University of Washington, Box 357730, Seattle, WA 98195, USA
Location/Qualifiers
FEATURES
source 1..5491
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.2"
<3..>108
/note="putative"
repeat_region /rpt_family="MTR"
/rpt_type=dispersed
73
/frequency="0.01"
/replace="t"
<207..>295
/note="putative"
repeat_region /rpt_family="MTR"
/rpt_type=dispersed
308
/frequency="0.01"
/replace="t"
<320..>339
/note="putative"
repeat_region /rpt_type=tandem
<340..>637
/note="putative"
repeat_region /rpt_family="Alu"
/rpt_type=dispersed
471
/frequency="0.01"
/replace="g"
545
/frequency="0.01"
/replace="t"
560
/frequency="0.22"
/replace="t"
624
/frequency="0.07"
/replace="c"
<638..>718
/note="putative"
repeat_region /rpt_type=tandem
832
/frequency="0.45"
/replace="t"
MRNA join(1060..1094,1855..1920,3013..3205,3786..4645)
/product="apolipoprotein-E"
1163
/frequency="0.35"
/replace="c"
1522
/frequency="0.01"
/replace="a"
1575
/frequency="0.01"
/replace="t"
join(1878..1920,3013..3205,3786..4503)
/note="APOE"
/codon_start=1

/product="apolipoprotein-E"
/protein_id="AAG27089.1"
/db_xref="GI:11034801"
/translation="MKVIMAAIIYFLAGCOAKVQAVETPEPELROOTEMQSGRM
ELALGRFMDYLRWVQTLSEQVEBELSSQVTOBLRANDETPELKLKXSELEBOLTP
VAELTRARSKELQAQAARLGDMEVCGLVQYRGEVQAMLGQSTELRVLASHLR
KLRLRLRDADLDQRLLAVYQAGAREGARGLSAJRERGLPLVEGRVRAAVGSLAG
QPLQERQAQMGGERLARMEKMSRTDRIDEVKEQVAVAEVRARLQEOAOOIRLQAEAFQ
ARLKSWEPEPLVEDMORONAGLVEKVOAAVGTSAAPVSPDNH"
1998
/frequency="0.10"
/replace="a"
<2124..>2435
/note="putative"
repeat_region /rpt_family="Alu"
/rpt_type=dispersed
2440
/frequency="0.21"
/replace="a"
<2569..>2848
/note="putative"
repeat_region /rpt_family="Alu"
/rpt_type=dispersed
2907
/frequency="0.01"
/replace="g"
3106
/frequency="0.01"
/replace="c"
<3472..>3588
/note="putative"
repeat_region /rpt_family="Alu"
/rpt_type=dispersed
3673
/frequency="0.01"
/replace="g"
3701~3702
/frequency="0.01"
/replace="ct"
3937
/frequency="0.14"
/replace="c"
4036
/frequency="0.01"
/replace="t"
4075
/frequency="0.07"
/replace="t"
repeat_region /rpt_type=tandem
<4755..>5056
/note="putative"
repeat_region /rpt_family="Alu"
/rpt_type=dispersed
5065..>5476
/note="putative"
repeat_region /rpt_family="Alu"
/rpt_type=dispersed
5229~5230
/frequency="0.03"
/replace="gg"
5229
/frequency="0.07"
/replace="t"
5229~5230
/frequency="0.40"
/replace="g"
5230
/frequency="0.13"
/replace="a"
5361
/frequency="0.06"
/replace="c"
BASE COUNT 1029 a 1663 c 1603 g 1196 t
ORIGIN

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:08:17 ; Search time 178.728 Seconds
(without alignments)
3994.237 Million cell updates/sec

Title: US-09-827-854-17

Perfect score: 1589
Sequence: 1 MKVLMAALVTFLLAGCCAKV.....VEKVOAAGTSAPVPSDNH 317

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09827854/Unat_11032003_101609_27476/app.query.fasta.1.3576
-DB=N.Geneseq_101002 -OFMT=fastp -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsm2 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptco -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09827854 -CGCN.1.1.1201 -Erunat_11032003_101609_27476 -NCPU=6 -ICPU=3
-NO.XLPXY -NO.MMAP -LARGEQUERY -NEG1.SCORER=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-GAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N.Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1589	100.0	1156	24	AAD22050	Human apolipoprotein
2	1582	99.6	1156	22	AAF84314	Human ApoE2 coding
3	1582	99.6	1156	24	AAD22049	Human apolipoprotein
4	1570	98.8	954	24	AAD26035	Human apolipoprotein
5	1570	98.8	1110	7	AA60409	Human apolipoprotein
6	1570	98.8	1147	22	ABA83113	Apolipoprotein E o
7	1570	98.8	1156	22	AAF84315	Human ApoE3 coding
8	1570	98.8	1156	24	AAD22048	Human apolipoprotein
9	1566	98.6	1156	24	AAD22052	Human apolipoprotein
10	1562	98.3	1156	24	AAD22051	Human apolipoprotein
11	1561	98.2	1157	17	AA706957	Human apolipoprotein
12	1561	98.2	1157	24	ABN95746	Gene #2244 used to
13	1561	98.2	1157	24	ABK64514	Human benign prost
14	1561	98.2	1157	24	ABL65450	Lung cancer relate
15	1560	98.2	1110	6	AA550450	Sequence encoding
16	1558	98.0	1136	22	AAF84316	Human ApoE4 coding
17	1558	98.0	1156	24	AAD22047	Human apolipoprotein
18	1466.5	92.3	1279	22	AA522437	Human cDNA encodin
19	1447.5	91.1	1107	19	AA757556	Human apolipoprotein
20	1381.5	86.9	9360	24	ABL31915	Human ApoE genomic
21	1381.5	86.9	10716	24	AAD26034	Human apolipoprotein
22	1376	86.6	3805	20	AA209524	Human Apo E genomic
23	1376	86.6	3805	20	AA209526	Human Apo E genomic
24	1361.5	85.7	10716	24	AAD26108	Human Apo E genomic
25	1148.5	72.3	965	24	AAD32081	Human alpha-1-anti
26	1148.5	72.3	5617	24	AAD32077	Human albumin prom
27	1148.5	72.3	6026	24	AAD32075	Nucleotide sequenc
28	1123	70.7	1126	19	AAV28159	ApoE4Lx2 protease
29	970	61.0	936	15	AAQ69101	Human ApoE4Lx2 CDN
30	970	61.0	936	17	AA718070	Partial human apol
31	969	61.0	660	18	AA769792	Human ApoE4 CDNA.
32	932	58.1	597	17	AA718068	ApoE4L protease cd
33	923	58.1	597	15	AA609099	Human cDNA encodin
34	869	54.7	1381	22	AA522673	Human cDNA associ
35	852.5	53.7	786	21	AAF18114	EST clone CP147.
36	663	41.7	600	20	AAV89595	Human cDNA used to
37	651	41.0	407	24	ABK34238	Colon adenocarcino
38	625	39.3	478	24	ABL62679	Thyroid cancer rel
39	625	39.3	478	24	ABL67340	Human diagnostic a
40	625	39.3	478	24	AA503049	Human secreted pro
41	594	37.4	499	22	AA503049	Human apolipoprote
42	504	31.7	405	21	AA503049	Human apolipoprote
43	499	31.4	330	12	AA503049	Human EST-derived
44	485	30.5	345	22	AA503049	Human secreted exp
45	410	25.8	382	21	AAA44191	Mouse secreted exp

ALIGNMENTS

RESULT 1
AAD22050 standard: DNA; 1156 BP.

12-FEB-2002 (first entry)

Human apolipoprotein E (apoE) allele, apoE1 DNA.

Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;

hypertiglyceridaemia; low density lipoprotein; LDL; ds.

Homo sapiens.

Key Location/Qualifiers
CDS 61..1014
/*tag= a


```

FT      sig-peptide      /product= "Human apoE allele, apoE1"
FT      mat-peptide      /*tag= b
FT      61..114
FT      115..1011
FT      /*tag= c
FT      /product= "Mature human apoE allele, apoE1"
XX
PN      WO200171136-A1.
XX
PD      18-OCT-2001.
XX
PF      06-APR-2001; 2001MO-US11358.
XX
PR      06-APR-2000; 2000US-0544386.
PR      04-OCT-2000; 2000US-0679088.
PR      05-APR-2001; 2001US-0827854.
XX
PA      (KOSP-) KOS PHARM INC.
PA      (UYBO-) UNIV BOSTON.
XX
PI      Zannits VI, Kypreos KE;
XX
DR      WPI: 2002-010885/01.
DR      P-PsDB; AAE13296.
XX
PT      New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT      cholesterol, delaying the onset of or treating atherosclerosis in
PT      mammal, without inducing hypertriglyceridemia
XX
XX      Claim 14; Page 82; 91pp; English.
XX
PS      The present sequence is a human apolipoprotein E (apoE)
CC      allele, apoE1 DNA. The apoE lipoproteins are useful for
CC      lowering cholesterol, delaying the onset of atherosclerosis,
CC      treating or regressing atherosclerosis without inducing
CC      hypertriglyceridemia, in a mammal lacking an endogenous,
CC      normally functioning apoE gene or low density lipoprotein (LDL)
CC      receptor or is at risk for developing atherosclerosis due to
CC      accumulation of lipoprotein remnants in the bloodstream or having
CC      a defect in remnant removal.
XX
SQ      Sequence 1156 BP; 209 A; 367 C; 431 G; 149 T; 0 other;
Alignment Scores:
Pred. No.: 2.23e-115 Length: 1156
Score: 1589.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0
US-09-827-854-17 (1-317) x AAD22050 (1-1156)
OY      1 MetLysValLeuTrpAlaIleuLeuValThrPheLeuAlaGlySerGlnAlaLysVal 20
DB      61 ATGAAGTTCTGTGGGCTGGTGTGTCGACATTCCTCGCAGAGTCCAGGCCAAGGTTG 120
OY      21 GlnGlnAlaValAlaGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnProGln 40
DB      121 GACCAAGCGGTGAGACAGACAGCCGAGCCGAGCTCGCCAGACAGACCGAGTGGCAGAGC 180
OY      41 GlnGlnAlaArgTrpGlnLeuAlaLeuGlnArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB      181 GGCACAGCGCTGGGAACTGCGACCTGGCTCTTTGGATTACTGCGTGGGTGAGACACA 240
OY      61 LeuSerGlnGlnValGlnGlnLeuLeuLeuSerSerGlnValThrGlnGlnLeuAlaGlnAla 80
DB      241 CTGTCTGACAGACAGTGAAGAGAGGCTGCTCAGCTCCAGAGTCAACCGAAGTGAAGGCGG 300
OY      81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnGlnLeu 100
DB      301 CTGATGAGACAGACCATGAAGAGATTGAAGCGCTACAAATCGGAATGAGAGACAACTG 360

```

```

OY      101 ThrProValAlaGlnGlnThrTrpArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
DB      361 ACCCGGTGGCGGAGAGAGACCGCGGCACGCTCTCCAGAGAGCTGCAGGCGCGCAGGCC 420
OY      121 ArgLeuGlnValAspMetGlnAspValLysGlnLysLeuValGlnThrArgGlyGlnVal 140
DB      421 CGGCTGGCGCGGACATGAGAGAGCTGTGCGCGCGCTGCTGAGTACACCGCGGAGGAG 480
OY      141 GlnAlaMetLeuAspGlnSerThrGlnGlnLeuLeuArgValArgLeuAlaSerHisLeuArg 160
DB      481 CAGGCCATGCTGCACACAGACACGAGAGAGCTCGGGGTGGCTCCGCTCCACCTGGCC 540
OY      161 LysLeuArgLysArgLeuLeuArgAspAlaAspLeuGlnLysCysLeuAlaValTyr 180
DB      541 AAGCTCGTAAAGCGGCTCCCGCAGATGCCGATGACCTCGAAGAGCTCGGACAGTGTAC 600
OY      181 GlnAlaGlnValAlaArgGlnGlnGlnValArgGlyLysSerAlaIleArgGlnArgLeuGln 200
DB      601 CAGCGCGGGCGCGGAGGGCGCGAGCGGCGGCTCAGCGCCATCCGCGAGCGCTGGGG 660
OY      201 ProLeuValGlnGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
DB      661 CCCCTGCTGGAAACAGGGCGCGCTGCGGGCGCGCCACTGTGGGCTCCCTGGCGCGCAGCGG 720
OY      221 LeuGlnGlnValArgAlaAlaThrGlyLysArgLeuArgAlaArgMetGlnGlnMetGly 240
DB      721 CTACAGAGACGGGCGCAGGCTGTGGGCGAGCGGCTCGCGCGCGATGAGAGATGGGCG 780
OY      241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
DB      781 AGCGGGACCGCGGACCGCGCTGCGAGGAGGAGACAGCTGGCGAGGTGGCGCGCAAG 840
OY      261 LeuGlnGlnAlaGlnGlnGlnLysArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB      841 CTGAGAGACAGAGCCACAGACAGATACGCTGCGAGCGGAGGCTTCAGGCGCGGCTCAAG 900
OY      281 SerTrpPheGlnProLeuValGlnAspMetGlnArgGlnThrPalaGlyLeuValGlnLys 300
DB      901 AGCTGTGTGAGAGCCCTGTGTGAAAGATGACAGCCAGTGGGCGCGGCTGTGGAGAG 960
OY      301 ValGlnAlaAlaValAlaGlnThrSerAlaIleProValProSerAspAsnHis 317
DB      961 GTGACAGCTGTGCGTGGCGACACGCGCGCCCTGTGCGCCAGGACATATCAC 1011
RESULT 2
AAF84314
ID      AAF84314 standard; cDNA; 1156 BP.
XX
AC      AAF84314;
XX
XX      21-JUN-2001 (first entry)
DE      Human ApoE2 coding sequence.
XX
XX      Human; ApoE2; Alzheimer's disease; arteriosclerosis; ss.
XX
OS      Homo sapiens.
XX
FH      Key Location/Qualifiers
FT      CDS 61..1014
FT      /product= "Human ApoE2"
XX
PD      JP2001017028-A.
XX
PF      23-JAN-2001.
XX
PR      28-APR-2000; 2000JP-0128919.
XX
PR      06-MAY-1999; 99JP-0125647.
XX
PA      (MITU ) MITSUBISHI CHEM CORP.
XX

```


DR WPI: 2001-285406/30.
 DR P-PSDB: AAB80996.
 XX
 PT New apoE humanized mammalian cell useful for screening for agents
 PT useful for treating or preventing Alzheimer's disease and
 PT arteriosclerosis -
 XX
 PS Disclosure: Page 11-12; 22pp; Japanese.
 XX
 CC The present invention relates to an ApoE humanised mammalian cell. The
 CC present sequence is the coding sequence for human ApoE2, which was used
 CC in the method of the present invention. The ApoE humanised mammalian cell
 CC can be used for screening for agents useful for treating or preventing
 CC Alzheimer's disease and arteriosclerosis.
 XX
 SQ Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;
 Alignment Scores:
 Pred. No.: 7.87e-115 Length: 1156
 Score: 1582.00 Matches: 316
 Percent Similarity: 99.68% Conservative: 0
 Best Local Similarity: 99.68% Mismatches: 1
 Query Match: 99.56% Indels: 0
 DB: Gaps: 0
 US-09-827-854-17 (1-317) x AAF84314 (1-1156)
 QY 1 MetLysValLeuThrPalaIalaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 61 ATGAAGGTTCTGTGGCTCGTGGCTGTGTCACATTCCTGGCAGATGCCAGGCCAAGGTG 120
 QY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGluThrGluThrGlnSer 40
 DB 121 GAGCAAGCGGTGAGACACAGCCGGAGCCGAGCTGCGCAGACAGCCAGTGGCAGAC 180
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 DB 181 GGCACACGCTGGGAACCTGGCAGCTGGCTGCTTTGGGATTAAGTGGCTGGGCTGGCAGACA 240
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB 241 CTGTGTGACAGCTGACAGAGAGAGCTGCTCACCAGTCAACCCAGAGACTGAGGGCG 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
 DB 301 CTGATGAGACAGACATGATGAAGACTTGAAGGCTTACAAATCGGAAGTGAAGAGACACTG 360
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 361 ACCCGGTGGCGAGAGAGACGGCGGCGCTGTCCAAAGAGCTGCAGGCGGCGCAGGCC 420
 QY 121 ArgLeuGluAlaAlaAspMetLysAspValLysGlyArgLeuValGluThrArgGlyGluVal 140
 DB 421 CGGCTGGGCGGACATGAGAGAGCTGTGGCGGCTGTGTGACAGTACCGCGCGAGGGTG 480
 QY 141 GlnAlaMetLeuAspGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 481 CAGGGCATGCTGGCGCAGAGACACCGAGGCTGCGGCTGCGCTCCACCTCGCGC 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 DB 541 AAGCTGCGTAAGCGGCTCTCCGCGATGCCATGACCTGCAGAAAGTGGCTGGCAGGTAC 600
 QY 181 GlnAlaGlyAlaArgGlnGluValGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 601 CAGGCGCGGGCGCGGAGGCGCCGAGCGCGCTCAGCGCATCCGACAGCGCGCTGGGG 660
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 DB 661 CCCCTGTGTGAACAGGGCGCGCTGGCGCGCCACTGTGTGGCTCCCTGGCGCGCAGCGG 720
 QY 221 LeuGlnGluArgAlaGlnAlaThrGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240

DB 721 CTACAGAGACGGGCGGAGGCTGTGGGCGAGCGGCTGTCCGCGCGGATGAGAGATGGGC 780
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 DB 781 ACCCGGACCCGGGACCGCTGTGACAGGTGAAGAGAGCAGGTGGAGGTGGCGCCGCAAG 840
 QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 DB 841 CTGGAGAGACAGGCCAGAGATACCGCTGCGAGGCCGAGGCTTCCAGGCCCGGCTCAAG 900
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpPalaGlyLeuValGluLys 300
 DB 901 ACCTGGTTGAGACCCCTGTGGAGACATGACAGCCGACATGGGCGGCTGTGGAGAG 960
 QY 301 ValGlnAlaAlaValGlyThrSerAlaIleProValProSerAspAsnHis 317
 DB 961 GTGCAAGCTGCCGTGGGACACAGCGCGCCCTGTGTCCAGGACACATAC 1011
 RESULT 3
 AAD22049
 ID AAD22049 standard; DNA; 1156 BP.
 AC AAD22049;
 XX
 DT 12-FEB-2002 (first entry)
 XX
 DE Human apolipoprotein E (apoE) isoprotein, apoE2 DNA.
 XX
 KW Human; apolipoprotein E; apoE; cholesterol; arteriosclerosis;
 KW hypertriglyceridaemia; low density lipoprotein; LDL; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 61..1014
 FT /tag= a
 FT /product= "Human apoE isoprotein, apoE2"
 FT sig_peptide 61..114
 FT /tag= b
 FT mat_peptide 115..1011
 FT /tag= c
 FT /product= "Mature human apoE isoprotein, apoE2"
 PN W0200177136-A1.
 XX
 PD 18-OCT-2001.
 XX
 PF 06-APR-2001; 2001WO-US11358.
 XX
 PR 06-APR-2000; 2000US-0544386.
 PR 04-OCT-2000; 2000US-0670088.
 PR 05-APR-2001; 2001US-0827854.
 XX
 PA (KOSP-) KOS PHARM INC.
 PA (UYBO-) UNIV BOSTON.
 XX
 PI Zannis VI, Kyriacs KE;
 XX
 DR WPI: 2002-010885/01.
 DR P-PSDB: AAE13295.
 XX
 PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
 PT cholesterol, delaying the onset of or treating arteriosclerosis in
 PT mammal, without inducing hypertriglyceridaemia -
 XX
 PS Claim 14; Page 81-82; 91pp; English.
 XX
 CC The present sequence is a human apolipoprotein E (apoE)
 CC isoprotein, apoE2 DNA. The apoE lipoproteins are useful for
 CC lowering cholesterol, delaying the onset of arteriosclerosis,
 CC treating or regressing arteriosclerosis without inducing
 CC hypertriglyceridaemia, in a mammal lacking an endogenous,
 CC normally functioning apoE gene or low density lipoprotein (LDL)

CC for haplotyping and/or genotyping the APOE gene. The haplotyping
 CC methods of the invention are useful for improving the efficacy and
 CC reliability of several steps in the discovery and development of
 CC drugs for treating diseases associated with APOE activity, e.g.
 CC familial dysbetalipoproteinemia, type III hyperlipoproteinemia,
 CC atherosclerosis, and Alzheimer's disease. They are useful to validate
 CC APOE as a candidate agent for treating a specific condition or disease
 CC predicted to be associated with APOE activity and in the design of
 CC clinical trials of candidate drugs for treating a specific condition
 CC or disease predicted to be associated with APOE activity. Genotyping
 CC or haplotyping methods are useful to screen for compounds targeting
 CC APOE to treat a specific condition or disease associated with APOE
 CC activity. The present sequence is a cDNA encoding human APOE protein.
 CC APOE gene is located on chromosome 19q13.2.

XX
 SO Sequence 954 BP; 168 A; 290 C; 377 G; 119 T; 0 other;

Alignment Scores:

Pred. No.:	5.48e-114	Length:	954
Score:	1570.00	Matches:	315
Percent Similarity:	99.37%	Conservative:	0
Best Local Similarity:	99.37%	Mismatches:	2
Query Match:	98.80%	Indels:	0
DB:	24	Gaps:	0

US-09-827-854-17 (1-317) x AAD26035 (1-954)

QY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 1 ATGAAGCTTGTGGGCTCGTGGTGTGACATTCTCGCAGAGTCCAGGCAAGGATG 60
 QY 21 GUGGlnAlaValAlaGlnThrGluProGluLeuArgGlnGlnThrGlnSer 40
 DB 61 GAGCAAGCGGTGGAGACAGACCGGAGCCGAGCTCGCCAGACAGCCAGTGGCAGAC 120
 QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 DB 121 GGCACACGCTGGGAATGACATGCTGCTTTGGATTACTGCTGGCTGGGTCAGACA 180
 QY 61 LeuSerGlnGlnValGlnGlnLeuLeuSerSerGlnValThrGlnLeuArgAla 80
 DB 181 CTGTCTGACAGGTGAGAGAGCGGCTGCTCCAGTCCAGTCCAGCAGACCTGAGGCGC 240
 QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTrpLysSerGlnLeuGlnGlnLeu 100
 DB 241 CTGATGAGACAGACCATGAGAGATTGAAGGCTTACAAATCGGAAGTGGAGACACTG 300
 QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
 DB 301 ACCCGGTGGGAGAGAGACGCGGCGCTGTCCAGAGAGCTGCAGCGGCGGCGGCGC 360
 QY 121 ArgLeuGlnAlaAspMetLysAspValLysGlyArgLeuValGlnThrArgGlyGlnVal 140
 DB 361 CGGCTGGGGCGACATGAGAGCGTGTGCGCGCTGTGTGACAGTACCGCGCGGCGGTG 420
 QY 141 GlnAlaMetLeuAspGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 421 CAGGCGATCTCGGCGGACAGACCGGAGCTGCGGGTCTGCTCCCTCCCACTGCGCC 480
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspLeuGlnLysCysLeuAlaValTyr 180
 DB 481 AAGCTCGTAAAGCGGCTCTCCGCGATGCCGATGACCTCAGAAAGCGCTGGCAGTGTAC 540
 QY 181 GlnAlaGlyAlaArgGlnGlnValAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
 DB 541 CAGGCGGGGCGGCGGAGGCGCGCGCTCAGCGCATCGCCAGCGCGCTCGGGG 600
 QY 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
 DB 601 CCCCTGTGTGAAACAGGCGCGCTGCGGCGCCACTGTGGCTCCTTGGCGGCGGCGG 660
 QY 221 LeuGlnGlnArgAlaGlnAlaTrpGlnGlnArgGlnArgGlnMetGlnGlnGly 240

DB 661 CTACAGAGCGGCGCGGCTGGGGCGAGCGGCTTGCCGCGCGATGAGAGATGGC 720
 QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
 DB 721 AGCGGACCGCGGACCGCTGGGAGGTGAAGAGACAGTGGAGGTGGCGCCAG 780
 QY 261 LeuGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 DB 781 CTGAGAGACAGCGCCAGAGATACGCCCTGCAGCGCGAGGCTTCCAGGCGCGCTCAG 840
 QY 281 SerTrpPheGlnProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
 DB 841 ACTGTGTGAGACCCCTGCTGGAAAGACATGACAGCCGCACTGGGCGCGCTGTGGAGAG 900
 QY 301 ValGlnAlaAlaValGlnThrSerAlaAlaProValProSerAspAsnHis 317
 DB 901 GTGACGCTGCGGTGGGACACAGCGCGCCCTGTGTCCAGCAGCAATACAC 951

RESULT 5

AAN60409
 ID AAN60409 standard; DNA; 1110 BP.

AC AAN60409;

DT 01-JAN-1980 (first entry)

DE Human apolipoprotein-E.

KW Apolipoprotein-E; hyperlipidemia; arteriosclerosis; ss.

OS Homo sapiens.

FT key location/Qualifiers

FT CDS 15..968

FT AU8547513 -A

PD 24-APR-1986.

PF 17-SEP-1985; 85AU-0047513.

PR 11-JUN-1985; 85JP-0126989.

PA (MITU) MITSUBISHI CHEM IND KK.

DR WPI; 1986-150217/24.

DR P-PSDB; AAP60507.

PT New DNA sequence coding for human apolipoprotein-E - and

PS expression vectors and transformed cells contg. it

PS Disclosure: Fig 2; 45pp; English.

CC The encoded protein is used to treat subjects who are deficient in
 CC apolipoprotein-E (or who produce abnormal forms of this molecule)
 CC and therefore are likely to suffer from hyperlipidemia, resulting in
 CC arteriosclerosis. It can also be used to raise antisera for
 CC detecting the protein deficiency or production of abnormal forms.

SO Sequence 1110 BP; 198 A; 354 C; 415 G; 143 T; 0 other;

Alignment Scores:

Pred. No.:	6.51e-114	Length:	1110
Score:	1570.00 <td>Matches:</td> <td>315</td>	Matches:	315
Percent Similarity:	99.37% <td>Conservative:</td> <td>0</td>	Conservative:	0
Best Local Similarity:	99.37% <td>Mismatches:</td> <td>2</td>	Mismatches:	2
Query Match:	98.80% <td>Indels:</td> <td>0</td>	Indels:	0
DB:	7	Gaps:	0

US-09-827-854-17 (1-317) x AAN60409 (1-1110)

QY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20

|||||
DB 15 ATGAAAGTCTGTCGGCTGCTGTCAGATTCCTGCGAGATGCCAGGCCAAGTGT 74
QY 21 Gtuglglalalalaglturhrgltuprogltuprogltuleuargltngltmhrgtlwtppglnser 40
DB 75 GAGCAAGCGGTGGAGAGAGCGGAGCCGAGCTGGCCAGCAGACCGAGTGGCAAGC 134
QY 41 GlyGlnArgrtpgluleualaleuGlyArpPheTrrpAspyrLeuArGrTPValGlnThr 60
DB 135 GGGCAGCGCTGGGAACTGGCACTGGGTGCTTTTGGGATTACCTGGCTGGGTGGCAACA 194
QY 61 LeuSerGltuglInValGltngltugluleuLeuSerSerGlnValThrGltngltuleuArGAla 80
DB 195 CTGTCTGAGAGAGGTGCAGGAGGAGCTGCTCAGCTCCAGGTCAACCCAGGAACGTGAGGGCG 254
QY 81 LeuMetAspGlturhMetLysGltuleuLysAlATyrLysSerGltuleuGltngltuglInLeu 100
DB 255 CTGTATGACAGACACCATGAAGAGAGTTGAAGGCTTACAAATCGGAATCTGGAGAACACTGTG 314
QY 101 ThrProValAlaGltuglturhArGAlaArGLeuSerLysGltuleuGlnAlaAlaGlnAla 120
DB 315 ACCCGGTGGCGAGAGAGCGGGGCGAGCTGTCCAAAGAGCTGCAGCGCGGCGAGGCC 374
QY 121 ArGLeuGltAlaAspMetGltuaSpValCysGlyArGLeuValGltntYrArGlyGltuVal 140
DB 375 CGGCTGGGCGCCAGACAGAGAGAGCTGTGGCGCGCTGTGCAGTACCCGCGGAGAGTGT 434
QY 141 GtAlaMetLeuAspGlnserThrGltngltuleuArGValArGLeuAlaSerHisLeuArg 160
DB 435 CAGGCCATGCTCGGCCAGACACCGAGAGACTGTGGTGGCTCCGCTCCCTCCACCTGCC 494
QY 161 LysLeuArGlyArGLeuLeuArGAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
DB 495 AAGCTGGTAAAGCGGCTCTCCGCGATGCCGATGACCTCGAGAAGCCCTGGCAGTGTAC 554
QY 181 GtAlaAGtAlaArGgtGltngltAlaGltuArgGlyLeuSerAlaAlaGltuArgLeuGly 200
DB 555 CAGGCGCGGGCCGCGAGGGCGCGGAGCGGCGCTCAGCCCATCCGCGAGCGCCTGGGG 614
QY 201 ProLeuValGltngltngltArGValArGAlaAlaThrValGltSerLeuAlaGltngltPro 220
DB 615 CCCCTGGTGAACAGGCGCCGCTGCGGGCGGCCACACTGTGGGCTCCCTGGCCGCGGACCG 674
QY 221 LeuGltngltuArGAlaGlnAlaTrrpGlyGltuArgLeuArGAlaArGMetGltngltMetGly 240
DB 675 CTACAGAGAGCGGGCCCGCTGGGGGAGCGGCTGCGCGCGGATGAGAGATGGCG 734
QY 241 SerArGrThrArGrAspArGLeuAspGltuValLysGltngltInValAlaGltuValArGAlaLys 260
DB 735 AGCGGGAGCCGCGACCGCTGGAGAGGTGAAGGAGAGGTGGCGGAGGTGGCGGCCAAG 794
QY 261 LeuGltngltInAlaGltngltInLLeArGLeuGlnAlaGltuAlaPheGlnAlaArGLeuLys 280
DB 795 CTGGAGAGAGAGGCCAGCAGATACGCTGCAGGCCGAGCCCTTCAGAGCCGCCCTCAAG 854
QY 281 SerTrrpPheGltuproLeuValGltuaSpMetGlnArGlnTrrpAlaGlyLeuValGltLys 300
DB 855 AGCTGGTTCCAGGCCCTCGTGGTGAAGACATCAGCGCCAGAGGGCGGCTGTGTGGAGAAG 914
QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 915 GTTCAGAGCTCCCTGGGCGACACGCGGCCCTGTGCGCCAGCACAATCTAC 965
RESULT 6
ABA83113
ID ABA83113 standard; DNA: 1147 BP.
XX
AC ABA83113;
XX
DT 08-FEB-2002 (first entry)
XX
DE Apolipoprotein E ovarian tumour marker gene, SEQ ID NO:63.
XX

KW Ovarian tumour marker gene; human; overexpression; upregulation;
KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;
KW identification; serous cystadenoma; borderline serous tumour;
KW serous cystadenocarcinoma; mucinous cystadenocarcinoma;
KW mucinous cystadenoma; borderline mucinous tumour; endometrioid carcinoma;
KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;
KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;
KW immune response pathway; cell proliferation regulation; protein folding;
KW membrane localised; secreted; therapeutic target; cytostatic;
KW gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN MO200175177-A2.
XX
PD 11-OCT-2001.
XX
PE 03-APR-2001; 2001MO-US10947.
XX
PR 03-APR-2000; 2000US-194336P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;
XX
DR WPI: 2001-626450/72.
XX
P-PSDB: ABB50287.
XX
PT Detecting and identifying ovarian tumor, identifying increased risk for
PT developing ovarian cancer, and determining effectiveness of ovarian
PT cancer treatment, by measuring expression level of ovarian tumor marker
PT gene .
XX
PS Claim 23; Page 105-106; 140pp; English.
XX
CC The invention relates to methods for diagnosing and prognosing ovarian
CC tumors in an individual via the detection and measurement of the
CC expression of ovarian tumour marker genes (ABA83081-ABA83122, ABA83180,
CC ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179,
CC ABA83181 and ABA83183). The methods of the invention are useful for
CC detecting an ovarian tumour in a patient, for identifying an individual
CC at increased risk for developing ovarian cancer, in prognostic tests for
CC assessing the relative severity of ovarian cancer, in tests for
CC monitoring a patient in remission from ovarian cancer and in tests for
CC monitoring disease status in a patient being treated for ovarian cancer.
CC The methods can additionally be used to identify a particular tumour as
CC being an ovarian tumour (i.e., an epithelial ovarian tumour selected from
CC serous cystadenoma, borderline serous tumour, serous cystadenocarcinoma,
CC mucinous cystadenoma, borderline mucinous tumour, mucinous
CC cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma,
CC clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner
CC tumour. The ovarian tumour marker genes of the invention were identified
CC using SAGE (serial analysis of gene expression) and were found to be
CC overexpressed in a broad variety of ovarian epithelial tumour cells
CC relative to normal ovarian epithelial cells. The marker genes are
CC implicated in immune response pathways, and the regulation of cell
CC proliferation and in protein folding, and many of these are membrane-
CC localised or secreted. In addition to their use as diagnostic and
CC prognostic markers, the ovarian tumour marker genes or their encoded
CC proteins may be used as therapeutic targets for the treatment and
CC prevention of ovarian cancer. Sequences ABA83081-ABA83122, ABA83180,
CC ABA83182 and ABA83184 represent the ovarian tumour marker genes of
CC the invention.
XX
SO Sequence 1147 BP; 210 A; 365 C; 425 G; 147 T; 0 other;

Alignment Scores:
Pred. No.: 6,75e-114 Length: 1147
Score: 1570.00 Matches: 315
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 98.80% Indels: 0
DB: Gaps: 22

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||
Db 421 CGGCTGGCGCCGACATGAGGAGACGTGTGGCGCCGCTGGTGCAGTACCCCGGAGAGTGTG 480
QY 141 GluAlaMetLeuAspGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||
Db 481 CAGGCCATGGCTCGGCGACGACCGAGAGACGTGGGTGGCTCCCTCCACCTGCGC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
|||
Db 541 AAGCTGGCTAAGGGCGCTCCCGGATGCGGATGACTGAGAAAGCCCTGGCAGTGTAC 600
QY 181 GluAlaGlyAlaArgGluGluGluGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||
Db 601 CAGGCCGCGGCGCCGAGGCGCCGAGCGGCGCTCAGCCCATCCCGAGCGCCTGGG 660
QY 201 ProLeuValGluGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
|||
Db 661 CCCCTGGTGGACAGGCGCCGCTGGGCGGCGCCACTGTGGCTCCCTGGCGGCGCAGCG 720
QY 221 LeuGlnGluArgAlaGlnAlaTyrPglyGluArgLeuArgAlaArgMetGluGluMetGly 240
|||
Db 721 CTACAGGACGCGGCGCCGAGCGCTGGCGGAGCGGCTGGCGCGGATGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAlaArgAlaLys 260
|||
Db 781 AGCGGACCGCGCGACCGCTGGAGCGAGGTAGAGACGAGGTGGCGGAGTGGCGCGCAG 840
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaIleArgLeuLys 280
|||
Db 841 CTGGAGAGAGAGCGCCCGACGATACGCTGCGAGCGCAGCGCTTCCAGCGCCCGCTCAAG 900
QY 281 SerTyrPheGluProLeuValGluAspMetGlnArgGlnTyrAlaGlyLeuValGluLys 300
|||
Db 901 AGCTGGTTCGAGCGCCCTGGTGGAAACATGACAGCGCCAGTGGCGGCGCTGGTGGAGAG 960
QY 301 ValGluAlaIleValGlyThrSerAlaIleProValProSerAspAsnHis 317
|||
Db 961 GTGCAGCTGCTCGTGGGACACGCGCCGCTGTGCCAGCAATCATC 1011
RESULT 8
AAD22048
ID AAD22048 standard; DNA: 1156 BP.
AC AAD22048;
XX
DT 12-FEB-2002 (first entry)
XX
DE Human apolipoprotein E (apoE) isoprotein, apoE3 DNA.
XX
KW Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;
XX hypertriglyceridaemia; low density lipoprotein; LDL; ds.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 61..1014
FT /product= "Human apoE isoprotein, apoE3"
FT /tag= a
FT sig_peptide 61..114
FT /tag= b
FT mat_peptide 115..1011
FT /tag= c
FT /product= "Mature human apoE isoprotein, apoE3"
XX
PN MO200177136-A1.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001MO-US11358.
XX
PR 06-APR-2000; 2000US-0544386.

PR 04-OCT-2000; 2000US-0679088.
PR 05-APR-2001; 2001US-0827854.
XX
XX (KOSP-) KOS PHARM INC.
PA (UYBO-) UNIV BOSTON.
XX
XX Zannis VI, Kypros KE;
PI
DR WPI: 2002-010885/01.
DR P-PSDB: AAEL3294.
XX
XX New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT cholesterol, delaying the onset of or treating atherosclerosis in
PT mammal, without inducing hypertriglyceridaemia -
XX
XX Claim 14: Page 81; 91pp; English.
CC The present sequence is a human apolipoprotein E (apoE)
CC isoprotein, apoE3 DNA. The apoE lipoproteins are useful for
CC lowering cholesterol, delaying the onset of atherosclerosis,
CC treating or regressing atherosclerosis without inducing
CC hypertriglyceridaemia, in a mammal lacking an endogenous,
CC normally functioning apoE gene or low density lipoprotein (LDL)
CC receptor or is at risk for developing atherosclerosis due to
CC accumulation of lipoprotein remnants in the bloodstream or having
CC a defect in remnant removal.
XX
XX Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 6, 81e-114 Length: 1156
Score: 1570.00 Matches: 315
Percent Similarity: 99.378 Conservative: 0
Best Local Similarity: 99.378 Mismatches: 2
Query Match: 98.80% Indels: 0
DB: Gaps: 0
US-09-827-854-17 (1-317) x AAD22048 (1-1156)
QY 1 MetLysValLeuTyrPheAlaIleLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||
Db 61 ATGAGAGTTCCTGTGGGCTGGCTGTGTGCATCTTCGACAGAGTCCAGGCGCAAGGTG 120
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTyrGlnSer 40
|||
Db 121 GAGCAAGCGGTGGAGACAGAGCGCGGAGCCGAGCTCGCCAGACAGCCAGTGGCAGAC 180
QY 41 GlyAlaArgTyrPheGluLeuAlaLeuGlyArgPheTyrAspTyrLeuArgTyrValGlnThr 60
|||
Db 181 GGCACGCGTGGGAACTGGGACTGGGTCCCTTTGGATTACCTGGCTGGGTGCAGACA 240
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValTyrGlnGluLeuArgAla 80
|||
Db 241 CTGTGACAGAGGTGACAGAGAGAGCTGTCAGCTCCAGGTACCCAGAGATGAGGGCG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
|||
Db 301 CTGATGACAGAGACCAATGGAAGGTGGAAGGCCCTACAAATCGAATCGAGGAACATG 360
QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
|||
Db 361 ACCCGGTGGCGAGAGACCGCGGACGCTGTCCAAAGAGCTGAGCGGCGCAGGCC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||
Db 421 CGGCTGGCGCGACATGAGGAGACGTGTGGCGCCGCTGGTGCAGTACCCCGGAGAGTGTG 480
QY 141 GluAlaMetLeuAspGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||
Db 481 CAGGCCATGCTGGCGACGACCGAGAGAGTGGCGGCTCCCTCCACCTGCGC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
|||

Db 541 AAGCTGCTAGCGGCTCTCCGCGATGACCTGACAGAGCGGCTGGCAGTGTAC 600
 QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db 601 CAGGCGGGGGCCCGGAGGGGCGCCAGCGGCTCAGGCGCATCCGCGAGCGCTGGGG 660
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
 Db 661 CCCCTGTGTGAACAGAGCGGCTGGGGCGGCGGCTGTGGCTCTGGCCGCGCAAGCCG 720
 QY 221 LeuGlnGluArgAlaGlnAlaThrPheGlyArgLeuArgAlaArgMetGluGluMetGly 240
 Db 721 CTACACGAGCGGCGGCGGAGCGCTGGGGCGGCGGCTGGCGGATGAGAGATGGGCG 780
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 Db 781 ACCCGGACCCGCGACCGCTGGAGCGAGTGAAGAGCAGAGTGGCGAGGTGGCGCCCAAG 840
 QY 261 LeuGlnGluGlnAlaGlnGlnIleLeuArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 Db 841 CTGGAGAGAGCAGAGCGGCGGAGATAGCGCTGGCGAGCGCGGAGCTTCCAGGCGCGCTCAAG 900
 QY 281 SerTrpPheGluProLeuValGluLysPheMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 Db 901 ACCTGCTTCGAGCCCTGTGTGAGAACATCCAGCCGCAAGTGGCGGCTGGTGGAGGAAG 960
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 961 GTGACAGGCTGCGGTGGGACACAGCGCGCCCTGTGCCAGGACATTCAC 1011
 RESULT 9
 AAD22052 standard; DNA; 1156 BP.
 ID AAD22052 standard; DNA; 1156 BP.
 AC AAD22052;
 XX 12-FEB-2002 (first entry)
 DT
 XX Human apolipoprotein E (apoE) allele, apoE2** DNA.
 DE Human apolipoprotein E (apoE) allele, apoE2** DNA.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 61..1014
 FT /tag= a
 FT /product= "Human apoE allele, apoE2**"
 FT sig_peptide 61..114
 FT /tag= b
 FT mat_peptide 115..1011
 FT /tag= c
 FT /product= "Mature human apoE allele, apoE2**"
 PN WO200177136-A1.
 PD 18-OCT-2001.
 XX
 PF 06-APR-2001; 2001WO-US11358.
 XX
 PR 06-APR-2000; 2000US-0544386.
 PR 04-OCT-2000; 2000US-0679088.
 PR 05-APR-2001; 2001US-0827854.
 XX
 PA (KOSP-) KOS PHARM INC.
 PA (UYBO-) UNIV BOSTON.
 XX
 PI Zannis VI, Kypros KE;
 DR WPI: 2002-010885/01.
 DR P-PSDB: AAE13298.
 XX

PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
 PT cholesterol, delaying the onset of or treating atherosclerosis in
 PT mammal, without inducing hypertriglyceridemia
 PS Claim 14; Page 83; 91pp; English.
 XX
 CC The present sequence is a human apolipoprotein E (apoE)
 CC allele, apoE2** DNA. The apoE lipoproteins are useful for
 CC lowering cholesterol, delaying the onset of atherosclerosis,
 CC treating or regressing atherosclerosis without inducing
 CC hypertriglyceridemia, in a mammal lacking an endogenous,
 CC normally functioning apoE gene or low density lipoprotein (LDL)
 CC receptor or is at risk for developing atherosclerosis due to
 CC accumulation of lipoprotein remnants in the bloodstream or having
 CC a defect in remnant removal.
 SO Sequence 1156 BP; 207 A; 369 C; 432 G; 148 T; 0 other.
 Alignment Scores:
 Pred. No.: 1..4e-113 Length: 1156
 Score: 1566.00 Matches: 314
 Percent Similarity: 99.37% Conservative: 1
 Best Local Similarity: 99.05% Mismatches: 2
 Query Match: 98.55% Indels: 0
 DB: 24 Gaps: 0
 US-09-827-854-17 (1-317) x AAD22052 (1-1156)
 QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 61 ATGAAGGTTCTGTGGGCTGCGCTGTGGTGCATTCCTGGCAGATGCCAGGCCAAGGTG 120
 QY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 Db 121 GAGCAAGCGGTGGAGACAGAGCGGCGGAGCTGCGGCGCACAGACGAGTGGAGAAC 180
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlnArgPheTrpPaspTrpLeuArgTrpValGlnThr 60
 Db 181 GGCCAGCGCTGGGAACCTGGCACTGGGTGCTTTTGGGATTAACCTGCGTGGGTGCAGACA 240
 QY 61 LeuSerGlnGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 Db 241 CTGTCTGAGCAGCTGACAGAGAGAGCTGCTCAGCTCCAGAGTCAACAGAACTGAGAGGGG 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluGlnGluGlnLeu 100
 Db 301 CTGATGAGCAGACCATTAAGAGATTGAAGCGCTCAAAATCGAAGTGGAGGAACTGAC 360
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 Db 361 ACCCGGTGGCGGAGAGACAGCGGGCGACGGCTGTCCAGAGACTGCAGCGCGCAGGCC 420
 QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 Db 421 CGGCTGGCGCGGACATGAGAGAGCTGTGCGGCGCCCTGTGTCACTACCGCGCGCAGGTG 480
 QY 141 GlnAlaMetLeuAspGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 481 CAGGCCATGCTGCGCCAGAGCACCGAGGAGCTGCGGGGTGGCTGCTCCCTCCACTCTGGC 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspPheGlnGlnLysCysLeuAlaValTyr 180
 Db 541 AAGCTGCTGACAGGCTCTCTCGCATGCGATGACTGACAGAACGCGCTGGCATGTGAC 600
 QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db 601 CAGGCGGGGGCCCGGAGGGGCGCGAGCGGCGCTCAGCGCATCCGAGAGCGCTGGGG 660
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
 Db 661 CCCCTGTGTGAACAGAGCGGCTGGGGCGGCGGCTGTGGCTCTGGCCGCGCAAGCCG 720
 QY 221 LeuGlnGluArgAlaGlnAlaThrPheGlyArgLeuArgAlaArgMetGluGluMetGly 240

|||||
Db 721 CTACAGAGAGGGCCAGAGCGCTGGGGGAGCGCGCTGGCGCGGATGGAGGATGGGC 780
QY 241 SerArgThrArgSPArgLeuAspGluValValValGluGlnValAlaGluValAlaArgAlaLys 260
Db 781 AGCCGGAGCCCGGACCCCTGGAGCAGAGTGAAGAGCAGAGTGGCGGAGGTGGCGCCAG 840
QY 261 LeuGluGluGlnAlaGlnAlaGlnIleArgLeuGlnAlaGluValAlaPheGlnAlaArgLeuLys 280
Db 841 CTGGAGAGAGAGCCAGCAGATACGCTTCGACGCGAGCGCTTCACAGGCCCGCTCAAG 900
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGluValGluLys 300
Db 901 AGCTGGTTCGAGCCCTGGTGGGAAGACATGCACCGCCAGTGGCGCGGTGTGGAGAG 960
QY 301 ValGlnAlaValAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTCCAGGCTCCCTGGGACACAGCGCCCGCTGTGCCACGACATCTAC 1011
RESULT 10
AAD22051
ID AAD22051 standard; DNA; 1156 BP.
XX
AC AAD22051:
XX
DT 12-FEB-2002 (first entry)
XX
DE Human apolipoprotein E (apoE) allele, apoE2* DNA.
XX
KW Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;
KW hypertriglyceridaemia; low density lipoprotein; LDL; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 61..1014
FT /tag= a
FT /product= "Human apoE allele, apoE2*"
FT sig_peptide 61..114
FT /tag= b
FT mat_peptide 115..1011
FT /tag= c
FT /product= "Mature human apoE allele, apoE2*"
XX
PN WO200177136-A1.
XX
PD 18-OCT-2001.
XX
PE 06-APR-2001; 2001WO-US11358.
XX
PR 06-APR-2000; 2000US-0544386.
PR 04-OCT-2000; 2000US-0679088.
PR 05-APR-2001; 2001US-0827854.
XX
PA (KOSP-) KOS PHARM INC.
PA (UYBO-) UNITV BOSTON.
XX
PI Zannis VI, Kypros KE;
XX
DR MPI: 2002-010885/01.
DR P-PSDB; AAE13297.
XX
XX
PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT cholesterol, delaying the onset of or treating atherosclerosis in
PT mammal, without inducing hypertriglyceridaemia -
XX
XX
PS Claim 14; Page 82; 91pp; English.
XX
CC The present sequence is a human apolipoprotein E (apoE)
CC allele, apoE2* DNA. The apoE lipoproteins are useful for
CC lowering cholesterol, delaying the onset of atherosclerosis,
CC treating or regressing atherosclerosis without inducing
CC hypertriglyceridaemia, in a mammal lacking an endogenous,

CC normally functioning apoE gene or low density lipoprotein (LDL)
CC receptor or is at risk for developing atherosclerosis due to
CC accumulation of lipoprotein remnants in the bloodstream or having
CC a defect in remnant removal.
XX

Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;

Alignment Scores:

Pred. No.:	2, 87e-113	Length:	1156
Score:	1562.00	Matches:	314
Percent Similarity:	99.05%	Conservative:	0
Best Local Similarity:	99.05%	Mismatches:	3
Query Match:	98.30%	Indels:	0
DB:	24	Gaps:	0

US-09-827-854-17 (1-317) x AAD22051 (1-1156)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAAGTTCTGTGGCTGCTGCTGTGTGCATTCCTTGGCAGATGCGAGCCAAAGTG 120
QY 21 GlnGlnAlaValGlnThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAAGCGGTGGAGACAGACCGGAGCCCGAGCTCGCCAGACAGCCAGTGGCAGAC 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlnArgPheTrpAspTrpLeuArgTrpValGlnThr 60
Db 181 GGCACAGCTGGGAGACTGGCACTGGGTCTTTGGATTAACCTGGGTGGGACAGACA 240
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGTCTGAGCAGGTGACAGAGAGGAGCTGCAGCTCCAGGTACCCAGAACTGAGAGGCG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu 100
Db 301 CTGATGAGACGACCATGAAGAGGTGAAGGCTTACAAATTCGAACTGAGACAACTG 360
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 361 ACCCGGTTGGCGGAGAGACCGGCGAGCGGTGTCCAAAGAGCTGACGCGCGCAGGCC 420
QY 121 ArgLeuGlnAlaAspMetLysPvalCysGlyArgLeuValGlnTrpArgGlyGluVal 140
Db 421 CGGCTGGCGCGGACATGAGAGGACGTGTGCGCGCTGCTGTCAGTACCGCGCGAGGTG 480
QY 141 GlnAlaMetLeuAspGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTGGCCAGACACCGAGAGCTGCGGCTCGCTCCCACTGCGCC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 541 AAGCTGTGAAGGGGCTCTCCGCGATGCGATGACCTGCAGAAAGCGCTCGCAGGTAC 600
QY 181 GlnAlaGlyAlaArgGluGlnAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 601 CAGCGCGGCGCCCGCAGAGCGCCGCGGCTCAGCCATCCGACAGCGCTCGGGG 660
QY 201 ProLeuValGluGlnGlnValArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGTGGAAACAGAGGCGCGCTGGCGGCGCCACATGTGGCTCCTGGCGCGCAGCGG 720
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
Db 721 CTACAGAGAGCGGGCCAGAGCTGGGGCGAGCGGTGCGCGCGAGTGAAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 781 AGCCGGAGCCCGGACCGCTGGAGCAGAGTGAAGAGACAGTGGCGAGTGGCGCGCAAG 840
QY 261 LeuGluGluGlnAlaGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 841 CTGGAGAGACAGCCAGCAGATACGCTGCAGGCGCGAGGCTTCAGGCGCGCTCAAG 900

QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluIlys 300
|||||
DB 901 AGCTGGTGGAGCCCTGGTGGAGAGACATGACGCCCATGGGCGGTGGTGGAGAG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
|||||
DB 961 GTGCAGGCTGCGGTGGGACACAGCGCGCCCTGTGCTCCAGGACATCAG 1011
RESULT 11
AAT06957
ID AAT06957 standard; cDNA to mRNA; 1157 BP.
XX
AC AAT06957;
DT 19-JUN-1996 (first entry)
XX
DE Human apolipoprotein-E (ApoE) cDNA.
XX
KW Recombinant; human; apolipoprotein-E; ApoE; insect cells; larva;
KW Manduca sexta; Autographa californica nuclear polyhedrosis virus;
KW haemolymph; lipid complex; biologically active; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc_feature 62..119
FT note="tag= a
FT note="misc_signal"
FT CDS 62..1015
FT tag= b
XX
PN US5472858-A.
XX
PD 05-DEC-1995.
XX
PE 04-JUN-1991; 91US-0709949.
XX
PR 04-JUN-1991; 91US-0709949.
XX
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Attie AD, Beckage NE, Gretsch DG, Sturley SL;
XX
DR WPI: 1996-029812/03.
XX
DR P-PSDB: AAR86791.
XX
PT Prod. of recombinant apo:lipoprotein E in insects - by infecting
PT Manduca sexta larvae with recombinant Autographa californica
PT nuclear polyhedrosis baculovirus vector.
XX
PS Disclosure: Columns 11-14; 10pp; English.
XX
CC Recombinant human apolipoprotein-E (ApoE) (AAR86791) can be produced
CC by preparing a genetic construct (contg. an ApoE-encoding sequence,
CC e.g. AAT06957, and flanking regulatory sequences enabling the protein
CC to be expressed in insect cells), which is then introduced into a
CC Manduca sexta larva (using a recombinant Autographa californica
CC nuclear polyhedrosis virus) and recovering the protein from the
CC haemolymph of the larval host. The ApoE produced is in a form
CC sufficiently complexed with lipids to be biologically active, which
CC cannot be achieved in insect cell cultures, and can therefore be
CC used in therapeutic applications.
XX
SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;

Alignment Scores:

Pred. No.: 3,44e-113 Length: 1157
Score: 1561.00 Matches: 313
Percent Similarity: 98.74% Conservative: 0
Best Local Similarity: 98.74% Mismatches: 4
Query Match: 98.24% Indels: 0
DB: 17 Gaps: 0

US-09-827-854-17 (1-317) x AAT06957 (1-1157)
QY 1 MetLysValLeuTrpAlaAlaLeuValTrpPheLeuAlaGlyLysGlnAlaIlysVal 20
|||||
DB 62 ATGAAGGTTCTGTGGGCTCGTTGCTGTGACATTCCTGGCAGAGATGCCAGGCCAAGG 121
QY 21 GlnGlnAlaValAlaGlyThrGluProGluProGluLeuArgGlnGlnThrGluProGlnSer 40
|||||
DB 122 GAGCAAGCGGTGGAGACACAGCCGAGCCCGAGCTCGCCAGCAGACCCAGTGGCAGAGC 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspPyrLeuArgTrpValGlnThr 60
|||||
DB 182 GCCAGCGCTGGGAACCTGACCTGCGCTCTTTGGATTACCTCGCTGGGTGGAGACA 241
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
DB 242 CTGTGTGACAGGTGAGAGAGAGCTGTGAGCTCCAGATCCCAAGTCAACCAAGACTGAGGGCG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
|||||
DB 302 CTGATGGAGAGACCATGAAGAGATTGAAGGCTTCAAAATCGGAACCTGAGAGACACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
|||||
DB 362 ACCCCGGTAGCGAGAGAGACGCGGCGCTGTCCAAAGAGCTGCAGACGCGCGAGGCC 421
QY 121 ArgLeuGlyAlaAspMetGluAspValLysGlyArgLeuValGlnThrArgGlyGluVal 140
|||||
DB 422 CGGCTGGGCGGACATGAGAGAGCTGTGCGGCTGTGTGATGATGACGGCGAGGTG 481
QY 141 GlnAlaMetLeuAspGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
DB 482 CAGGCCATGCTGGGCCAGAGACACGAGAGCTGCGGCTGCGCTGCCCTCCACCTGGCGC 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
|||||
DB 542 AAGCTGCTGAAGCGGCTCTCCGCGATCCGATGACTCCAGAAAGCGCTGGCAGGTAC 601
QY 181 GlnAlaGlyAlaArgGlnGluValGlnArgGlyLeuSerAlaAlaArgGluLysArgLeuGly 200
|||||
DB 602 CAGGCGGGGCGCGGAGGCGCGGCGGCGGCTGAGGCCATCCGCGAGCGCTGGG 661
QY 201 ProLeuValGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
|||||
DB 662 CCCCTGTGTGACAGAGGCGCGGCTGCGGCGCCACACTGTGGCTCCTGGCGCGCCAGCGG 721
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyLysArgLeuArgAlaArgMetGlnGluMetGly 240
|||||
DB 722 CTACAGAGACCGGGCCAGGCTGTGGGCGAGCGGCTGCGCGCATGTGAGAGATGGGC 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaIlys 260
|||||
DB 782 ACTCGGACCGCGGACCGCTGGAGCGAGTGAAGAGACAGAGTGGAGGTGGCGCCAGAG 841
QY 261 LeuGlnGluGlnAlaGlnGlnLysArgLeuGlnAlaAlaPheGlnAlaArgLeuLys 280
|||||
DB 842 CTGGAGAGACAGCCAGAGATACCTGCGAGCGCGAGAGCTTCCAGGCGCGCTCAAG 901
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluIlys 300
|||||
DB 902 ACCTGGTGTGAGCCCTGTGTGAAGACATGACAGCGCATGTGGCGGTGTGTGGAGAG 961
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
|||||
DB 962 GTGCAGGCTGCGGTGGGACACAGCGCGCCCTGTGCTCCAGGACATCAG 1012
RESULT 12
ABN95746
ID ABN95746 standard; DNA; 1157 BP.
XX
AC ABN95746;
XX
DT 13-AUG-2002 (first entry)

PA (NISR) JAPAN TOBACCO INC.
XX Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
XX WPI; 2002-257476/30.
XX
XX Identifying drugs for and diagnosing benign prostatic hyperplasia, by
PT detecting expression levels of one or more genes in prostate cells from
PT patient that are differentially regulated compared to normal prostate
XX cells -
XX
PS Disclosure: Page 239-240; 444pp; English.
XX
XX The invention relates to a method of diagnosing (I) the onset or
CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
CC or identifying an agent that modulates the onset or progression of BPH.
CC The method is based on changes in gene expression in BPH tissue isolated
CC from patients exhibiting different clinical states of prostate
CC hyperplasia as compared to normal prostate tissue. (I) comprises
CC detecting the expression levels of one or more genes in prostate cells
CC from the subject that are differentially regulated compared to normal
CC prostate cells. (II) comprises preparing a first gene expression profile
CC of BPH cells or BPH-like cell population, exposing the cells to the
CC agent, preparing a second gene expression profile of the agent exposed
CC cells, and comparing the first and second gene expression profiles.
CC (I) is useful for diagnosing the onset or progression of BPH. (II) is
CC useful for identifying an agent that modulates the onset or progression
CC of BPH. The methods are useful to present information identifying
CC the expression level in a tissue or cells, by comparing the expression
CC level of genes given in the specification in the tissue or cells to the
CC level of expression of at least one gene in the database, and displaying the
CC expression levels of at least one gene in the tissue or cell sample
CC compared to the expression level in BPH. Agents using (II) are useful for
CC treating BPH or prostate cancer. ABK64106-ABK64860 represent human
CC benign prostatic hyperplasia gene sequences of the invention.
XX
XX Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other:
SQ
Alignment Scores:
Pred. No.: 3 44e-113
Score: 1561.00 Length: 1157
Percent Similarity: 98.74% Matches: 313
Best Local Similarity: 98.74% Mismatches: 4
Query Match: 98.24% Indels: 0
DB: 24 Gaps: 0
US-09-827-854-17 (1-317) x ABK64514 (1-1157)
OY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 62 ATGAAGGTTCTGGGCTCGCTGCTGTCACATTCCTGGCAGAGATGCCAGGCAAGT 121
OY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGlnGlnSer 40
DB 122 GAGCAGACGCGTGGAGACAGACGCGGAGCCGAGCTGCGCAGACAGACGAGTGGCAGAGC 181
OY 41 GLYGLAArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 182 GGCACAGCGCTGGGAATGGAGCTGGCTCTTTGGATTACTGTGGCTGGGAGACA 241
OY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuAlaGala 80
DB 242 CTGTCTGACAGAGTGCAGAGAGAGCTGCTCAGTCCCAAGTCAACCAAGTCAAGAGGCG 301
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnGlnLeu 100
DB 302 CTGATGAGAGACCATGTAAGAGTGAAGGCCCTACAAATCGGAACCTGAGAGAACACTG 361
OY 101 ThrProValAlaGlnGlnGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 362 ACCCCGCTAGCGAGGAGAGCGGGGCGACGCTGTCCAAAGAGCTGACGCGGCGGCGCC 421
OY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140

DB 422 CGGCTGGCGCGGACATGAGAGACGTGTGGCGGCGCTGTGTGCATGCCGCGGAGGTG 481
OY 141 GlnAlaMetLeuAspGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 482 CAGCGCATGCTCGCGCCAGAGACCGAGAGCTCGGGGTGGCTCCCTCCACCTGCGC 541
OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspLeuGlnLysCysLeuAlaValTyr 180
DB 542 AACCTCGTAAGCGGCTCTCCCGCATCCGATGACCTGCAGAAAGCGCTGGCAGTGTAC 601
OY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 602 CAGCGCGGGGCGCGCGAGGCGCGCGAGCGGCGCTCAGCGCCATCCGCGAGCGCTCGGG 661
OY 201 ProLeuValGluGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
DB 662 CCCCTGTGTGAGACAGGGCGCGCTGCGGCGCGCCACTGTGGGCTCCCTGGCGCGCAGCG 721
OY 221 LeuGlnGluArgAlaGlnAlaTrpGlyArgLeuArgAlaArgMetGluGlnMetGly 240
DB 722 CTACAGGAGCGGCGCCAGAGCTGGGCGAGCGCTGCGCGCGCGATGAGAGATGGG 781
OY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGlnValArgAlaLys 260
DB 782 AGTCGAGACCGCGACCGCTGCGAGAGGTGAAGAGAGAGTGGCGGAGTGGCGCGCAAG 841
OY 261 LeuGlnGluGlnAlaGlnGlnLeuArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 842 CTGGAGAGACAGCGCCAGCAGATACCTGACAGCGGAGGCTTCCAGGCGCCCTTCAG 901
OY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlnLeuValGlnLys 300
DB 902 AGCTGTGTCAGCGCTGCTGGTGAACATGCACGCCAGTGGCGCGGCTGGTGGAGAG 961
OY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 962 GTGCAGGCTGCGCTGGGACACACGCGCGCCCTGTGCTCCAGCAGCATATC 1012
RESULT 14
ABL65450
ID ABL65450 standard; DNA; 1157 BP.
XX
AC ABL65450;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:3787.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms tumor; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.

PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237315P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237586P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 (AVAL-) AVALON PHARM.
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX WPI; 2002-188264/24.

XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 PS Claim 1; SEQ ID 3787; 44pp; English.

XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL6164
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.

XX Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;

Alignment Scores:

Pred. No.: 3 44e-113 Length: 1157
 Score: 1561.00 Matches: 313
 Percent Similarity: 98.74% Conservative: 0
 Best Local Similarity: 98.74% Mismatches: 4
 Query Match: 98.24% Indels: 0
 DB: 24 Gaps: 0

US-09-827-854-17 (1-317) x ABL65450 (1-1157)

QY 1 MetlyValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
 Db 62 ATGAGGTTCTGTGGCTGGCTGTGTCACATTCCTGGCAGAGATCCAGGCCAAGGTG 121
 QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnInThrGluThrPglInser 40
 Db 122 GAGCAACGGCTGGAGACAGACGCGGAGCCGACCTGGCGCAGACACCGAGTGGCAGAGC 181
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheThrAspTyrLeuArgTrpValGlnThr 60
 Db 182 GCCCAGCGCTGGACAGCTGGCTGGCTTTGGGATTACCTGGCTGGCTGGCAGACA 241
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnLeuLeuArgAla 80
 Db 242 CTGCTGACGAGGTGGCAGGAGGAGCTGCTGACGCTCCCAAGTCAACCCAACTGAGGGCG 301
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluLeu 100
 Db 302 CTGATGGACAGACCATGAAAGGAGTTGAAAGCCCTACAAATCGGAAGTGGAGAAACTG 361
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluGlnAlaAlaGlnAla 120
 Db 362 ACCCGGCTAGCGAGGAGACCGCGGACGCGCTCTCCAGAGCTGCCAGAGCGCGCAGGCC 421
 QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnThrArgGlyGluVal 140
 Db 422 CGGCTGGCGCGGACATGAGGAGAGCTGTGGCGCGCTGGTGCAGTACCGCGCGAGAGTG 481
 QY 141 GlnAlaMetLeuAspGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 482 CAGCGCATGCTCGCGCAGACACCGAGAGCTCGGGTGGCGCTCCGCTCCGCTCCGCTCGC 541
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspLeuGlnLysCysLeuAlaValTyr 180
 Db 542 AAGCTGGTAAAGCGGCTCTCCGCGATCCCGATGACCTCGAGAAAGCGCTCGCAGGTAC 601
 QY 181 GlnAlaGlyAlaArgGluGluGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db 602 CAGCGCGGCG 661
 QY 201 ProLeuValGluGluGluArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 662 CCCCTGGTGAACAGGCG 721
 QY 221 LeuGlnGluArgAlaGlnAlaThrPglGlyArgLeuArgAlaArgMetGluMetGly 240
 Db 722 CTACAGAGAGCGGCG 781
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 Db 782 AGTCGACCG 841
 QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 Db 842 CTGGAGAGAGAGCG 901
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnThrAlaGlyLeuValGluLys 300
 Db 902 AGCTGGTTCAGAGCCCTGTGGAGAACATGACCGCCAGTGGCGCGCGCGCGCGCGCG 961
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 962 GTGACAGGTGCGCTGGGACACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1012

RESULT 15
 AANS0450
 ID AANS0450 standard; DNA; 1110 BP.
 XX
 AC AANS0450;
 XX
 DT 09-JAN-1992 (first entry)


```

XX DE Sequence encoding human apolipoprotein E.
XX KW Hyperlipaemia; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 15..968 /*tag= a
XX FT mat_peptide 69..965
XX FT /*tag= b
XX JP60118189-A.
XX PD 25-JUN-1985.
XX PF 29-NOV-1983; 83JP-0224980.
XX PR 29-NOV-1983; 83JP-0224980.
XX PA (MITU ) MITSUBISHI CHEM IND KK.
XX DR WPI: 1985-188003/31.
XX DR P-PSDB; AAP51204.
XX PT DNA fragment - contg. DNA which codes human apolipoprotein E for
XX PT treatment of hyperlipaemia.
XX PS Claim 3; Page 484; 8pp; Japanese.
XX CC The sequence may be used to produce the apolipoprotein E, useful in the
XX CC treatment of hyperlipaemia.
XX SQ Sequence 1110 BP; 198 A; 355 C; 414 G; 143 T; 0 other:

Alignment Scores:
Pred. No.: 3,93e-113 Length: 1110
Score: 1560.00 Matches: 313
Percent Similarity: 98.74% Conservative: 0
Best Local Similarity: 98.74% Mismatches: 4
Query Match: 98.17% Indels: 0
DB: 6 Gaps: 0

US-09-827-854-17 (1-317) x AAN50450 (1-1110)
QY 1 MetIysValLeuTrpAlaIalaLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20
DB 15 ATGAAGGTTCTGTGGCTCTGCTGTCACATTCCTGGCAGATGCCAAGGTG 74
QY 21 GluGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 75 GAGCAAGCGGTGGAGACAGCCGGAGCCGAGCTGCGCACAGACCGAGTGGCAGAGC 134
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 135 GGCAGAGCGCTGGGAACCTGACATGGGTCTTTTGGGATTACCTGGCGTGGCAGACA 194
QY 61 LeuSerGluGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 195 CTGTCTGACACGTGAGAGGAGGTGCTCAGCTCCAGGTCAACCGAGAACTGAGGGCG 254
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100
DB 255 CTGATGGACGAGACCATGTAAGAGTTGAGGCTTACAAATCCGAACCTGGAGAGACACTG 314
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaIaGlnAla 120
DB 315 ACCCGGTGGCGAGAGACGCGGACGCTGTCCAAGAGACTGACAGCGCGCCAGGCC 374
QY 121 ArgLeuGlyAlaAspMetLysPylCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 375 CGGCTGGCGCGACATGAGAGACGTGCGCGCTGTGTGACGTACCGCGCGAGGTG 434

```

```

QY 141 GlnAlaMetLeuAspGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 435 CAGGCCATGCTCGGCCAGACACCGAGAGCTGCGGGTCTGCGCTCCACCTCGCCG 494
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
DB 495 AAGCTGCTGTAACGGCTCTCCGCGATGCGATGACCTCCAGAAAGCGCTGCGAGTGC 554
QY 181 GlnAlaGlyAlaArgGluGlnLysAlaGluArgGlyLeuSerAlaThrArgGluArgLeuGly 200
DB 555 CAGCGCGGGCGCGGAGGGCGCGAGCGGCTCTCAGCGCATCCGCGAGCGCGCTGGGG 614
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 615 CCCCTGTGTGAACAGAGCGCGCGCTGCGCGCGCCACTGTGGCTCTCTGGCCGCGCAGCCG 674
QY 221 LeuGlnGluArgAlaGlnAlaThrPglGlyGluArgLeuArgAlaArgMetGluMetGly 240
DB 675 CTACAGAGAGCGGGCCAGGCTGTGGGCGAGCGGCTGCGCGCGATGAGAGAGATGGGC 734
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 735 ACCCGGACCCGCGACCGCTTGACAGAGTGAAGAGACAGTGGCGAGGTGGCGCCGCAAG 794
QY 261 LeuGluGluGlnAlaGlnGlnIleArgGluGlnAlaGluAlaPheGlnAlaArgLeuLys 280
DB 795 CTGGAGAGACAGGCCCGACAGATACGCTGACAGGCCGAGGCCCTTCCAGGCCCGCTCAAG 854
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 855 ACCTGTTGTGAGCCCTGTGTGAAGACATGACAGCGCCAGTGGCGCGGCTGTGGAGAG 914
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 915 GTGCAAGCTGCGGTGGGACCAAGCGCGCCCTGTGCCAGAGACATATAC 965

```

Search completed: March 14, 2003, 12:43:15
 Job time : 183.728 secs


```

: NAME/KEY: CDS
: LOCATION: 51..1001
: US-08-949-155-5

Alignment Scores:
Pred. No.: 5,98e-101
Score: 1123.00
Percent Similarity: 83.12%
Best Local Similarity: 69.38%
Query Match: 70.67%
DB: 4

Length: 1126
Matches: 222
Conservative: 44
Mismatch: 46
Indels: 8
Gaps: 3

US-09-827-854-17 (1-317) x US-08-949-155-5 (1-1126)
QY 1 MetLysValLeuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 51 ATGAGGTTCTGTGGTCTTTGGTGTACCTCTCGCAGATGCCGACAGAGAC 110

QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 111 GAGCCGGG-----CCGCCCGGAGGTGCACGTGTGTGGAGAGAGCCCAAG 158

QY 38 TrpGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrp 57
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 159 TGGCAGGAGCCAGCCCTGGAGCAGCCCTGGCCCTCTGGGATTACCTGGCCTGG 218

QY 58 ValGlnThrLeuSerGlnGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGlu 77
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 219 GTGCACCTCTGTCTACCAAGTGCAGAGAGAGCTCTACACCAAGTCCACCCAGAA 278

QY 78 LeuArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlu 97
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 279 CTGACGAGCTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 338

QY 98 GluGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAla 117
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 339 GCGCAGCTGGCCCTGGACCCGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 398

QY 118 AlaGlnAlaArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArg 137
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 399 GCGCAGGCGCCCTGGGCGCCGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 458

QY 138 GlyGluValGlnAlaMetLysAspGlnSerThrGlnGluLeuArgValArgLeuAlaSer 157
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 459 ACGAGAGTGCACAACTGTGGCGCCAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 518

QY 158 HisLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeu 177
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 519 CACCTGCGCAAGCTGCGCAGAGGCGCTGCCGACACCGAGAGAGAGAGAGAGAGAG 578

QY 178 AlaValTyrGlnAlaGlyAlaArgGluGlnGlyAlaArgGlyLeuSerAlaIleArgGlu 197
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 579 GCGGTGTACAGGCGGCGGCTGCCGCGAGGCGCCGAGCGAGCGTGTGCCCTCCCGAG 638

QY 198 ArgLeuGlyProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAla 217
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 639 CGCCTGGGCGCCCTGGTGTGAGCAGGCGCGATGCGCGCCGACCTGTAGTACAGAGCC 698

QY 218 GlyGlnProLeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGlu 237
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 699 GCGCAGCCCTCTGTGAGGCGCGCGAGAGCTGTGGCGCCCAACACTGTGGCGAGCG 758

QY 238 GluMetGlySerArgThrArgAspArgLeuAspGluValGlyGlnGlnValAlaGluVal 257
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 759 GAGATGGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 818

QY 258 ArgAlaLysLeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaLysLeuAla 277
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 819 CCGACCAAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 878

QY 278 ArgLeuLysSerTrpPheLeuProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeu 297
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 879 CTCCTCAAGGCTGTTGAGGCTCTGTGTGAAGACATACAGCGCGCAGTGGCGCGGCTG 938

```

```

QY 298 ValGluLysValGlnAlaValGly----ThrSerAlaAlaProValProSerAspAsn 316
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 939 GTGGAGAGATGACGTGGCGCGTGGAGCATRAGCTCTCCACCTCTGCGCCAGTATAT 998

RESULT 3
US-09-819-964-5
: Sequence 5, Application US/09819964
: Patent No. 6369294
: GENERAL INFORMATION:
:   APPLICANT: Piedrahita, Jorge A
:   TITLE OF INVENTION: Compositions and Methods for the
:                       Generation of Transgenic Animal Species
:   NUMBER OF SEQUENCES: 51
:   CORRESPONDENCE ADDRESS:
:   ADDRESSER: ARNOLD, WHITE AND DURKEE
:   STREET: P.O. Box 4433
:   CITY: Houston
:   STATE: TX
:   COUNTRY: US
:   ZIP: 77210-4433
:   COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk
:   OPERATING SYSTEM: IBM PC compatible
:   SOFTWARE: PatentIn Release #1.0, Version #1.30
:   CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/09/819,964
:   FILING DATE: 28-Mar-2001
:   CLASSIFICATION: <Unknown>
:   PRIOR APPLICATION DATA:
:   APPLICATION NUMBER: 08/949,155
:   FILING DATE: 09-MAY-1997
:   APPLICATION NUMBER: US 60/046,094
:   ATTORNEY/AGENT INFORMATION:
:   NAME: Hibler, David W.
:   REGISTRATION NUMBER: 41,071
:   TELEPHONE: (512) 418-3000
:   TELECOMMUNICATION INFORMATION:
:   TELEPHONE: (512) 418-3000
:   TELEFAX: (713) 789-2679
:   INFORMATION FOR SEQ ID NO: 5:
:   SEQUENCE CHARACTERISTICS:
:   LENGTH: 1126 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
:   FEATURE:
:   NAME/KEY: CDS
:   LOCATION: 51..1001
:   SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-819-964-5

Alignment Scores:
Pred. No.: 5,98e-101
Score: 1123.00
Percent Similarity: 83.12%
Best Local Similarity: 69.38%
Query Match: 70.67%
DB: 4

Length: 1126
Matches: 222
Conservative: 44
Mismatch: 46
Indels: 8
Gaps: 3

US-09-827-854-17 (1-317) x US-09-819-964-5 (1-1126)
QY 1 MetLysValLeuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 51 ATGAGGTTCTGTGGTCTTTGGTGTACCTCTCGCAGATGCCGACAGAGAGAC 110

QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 111 GAGCCGGG-----CCGCCCGGAGGTGCACGTGTGTGGAGAGAGAGAGAGAGAG 158

```



```

QY 38 TrrpGlnSerGlyGlnAtrpGlnLeuAlaLeuGlyAtrpPheTrpAspTyrLeuAtrp 57
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 159 TGGCAGGAGGAGCCGCTGGAGAGGAGCCCTGGCGCTTCTGGAGTATACCTGGCTGG 218
QY 58 ValGlnThrLeuSerGlnGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGln 77
    ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 219 GTCCAGTCCCTGTCTGACCAAGTGCAGAGGAGCTGCTGACCAAGGTCACCCAGAA 278
QY 78 LeuArgAlaLeuMetAspGlnThrMetGlyGlnLeuGlyAlaTyrLeuSerGlnLeu 97
    ||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 279 CTGACGAGAGCTGATAGGAGAGCATGACAGAGGTGAAGGCTTACCCGAGAGAGTGA 338
QY 98 GluGlnLeuThrProValAlaGlnGlnThrArgAlaArgLeuSerGlyLeuGlnAla 117
    ||||| ||||| ||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 339 GCCGAGCTGGGCCCCGTCAGCCAGAGAGACCCAGCGGCTGTCCAAAGAGAGCTGACGG 398
QY 118 AlagAlaAlaArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArg 137
    ||||| ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 399 GCCGAGGCCCCGCTGGGCGGACATGAGAGAGTGGCGCAACCGCTTGGTCTTACCGC 458
QY 138 GlyGlnValGlnAlaMetLeuAspGlnSerThrGlnGlnLeuArgValAlaArgLeuAlaSer 157
    ||||| ||||| ||||| ||||| :||| :||| :||| :||| :||| :|||
Db 459 AGGAGAGTGCACAACATGTTGGGCGACAGCACCGAGAGAGCTGGGAGCGCGCTGTCC 518
QY 158 HisLeuArgGlyLeuArgGlyArgLeuLeuArgAspAlaAspAspLeuGlnGlyCysLeu 177
    ||||| ||||| ||||| ||||| ||||| :||| :||| :||| :||| :|||
Db 519 CACCTCGCGCAAGCTGGCGCAAGCGGCTGCTCCGCGACACCGAGACCGCGAAGCGCGCT 578
QY 178 AlaValTyrGlnAlaGlyAlaArgGlnGlyAlaGlnArgGlyLeuSerAlaIleArgGln 197
    ||||| ||||| ||||| ||||| ||||| ||||| :||| :||| :||| :|||
Db 579 GCCGTGTACAGGCGGGGCGTGGCGAGGCGCGGCGAGCGGCTGACCGCTCCGCGAG 638
QY 198 ArgLeuGlyProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAla 217
    ||||| ||||| ||||| ||||| ||||| ||||| :||| :||| :||| :|||
Db 639 CGCCTCGGCGCCTGCTGGAGAGAGCGGCGCATTCGCGCCGCCACCTCGATACAGGCGC 698
QY 218 GlyGlnProLeuGlnGlnArgAlaGlnAlaTrrpGlyGlnArgLeuArgAlaArgMetGln 237
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| :||| :|||
Db 699 GGGCAGCGCGCTGGCGCGGCGCGGGAAGCTGGGGCGCGCAAGCTGCGCGAGCTGGAG 758
QY 238 GluMetGlySerArgTrpArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnVal 257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :|||
Db 759 GAGATGGGAGCGGAGCGGCGGAGCGGCGGCTGTGATGATGCTGAGAGAGAGAGTGG 818
QY 258 ArgAlaLeuSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 277
    ||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 819 CGCACCAGAAAGTGGAGAGAGAGAGAGAGAGTTCGCGCTGCAAGCGCGAGGATTCACAGCC 878
QY 278 ArgLeuLeuSerTrpPheGlnProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeu 297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :|||
Db 879 CTCTCTCAAAAGCTGTGTGAGCTGTGTGGAAGACATACAGGCGCGAGTGGCGCGCTG 938
QY 298 ValGlnThrLeuSerGlnGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 316
    ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 939 GTGAGAGAGATGACGTGGCGCTGAGCATTAAGCTCTCCACCTCTGCGCGCAGTATAT 998

```

RESULT 4
US-08-949-155-51
Sequence 51, Application US/08949155
Patent No. 6271436

GENERAL INFORMATION:
APPLICANT: Piedrahita, Jorge A
APPLICANT: Bazer, Fuller W
TITLE OF INVENTION: Compositions and Methods for the
TITLE OF INVENTION: Generation of Transgenic Animal Species
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE AND DURKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/949,155
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/027,338
; FILING DATE: 11-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,094
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hibler, David W.
; REGISTRATION NUMBER: 41,071
; REFERENCE/DOCKET NUMBER: TANK:177
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ. ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4267 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-949-155-51

```

US-09-827-854-17 (1-317) x US-08-949-155-51 (1-4267)

```

QY 1 MetLeuValLeuThrPheAlaIleLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    ||| ||| ||| ||||| ||||| :|||
Db 2448 ATGCTGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2489
QY 21 GluGlnAlaValGlnThrGlnProGlnProGlnLeuArg-----GlnGlnThrGln 37
    ||| ||| ||||| :||| :||| :||| :||| :||| :|||
Db 2490 GAGCCCGGG-----CCGCGCGGAGAGTGCAGCTGTGTGGAGAGAGCCAG 2537
QY 38 TrrpGlnSerGlyGlnAtrpGlnLeuAlaLeuGlyAtrpPheTrpAspTyrLeuAtrp 57
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2538 TGGCAGGAGGAGCCGCTGGGAGAGAGGCGCTGGCGCTTCTGGGATTAACCTGCGCTGG 2597
QY 58 ValGlnThrLeuSerGlnGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 77
    ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 2598 GTGCAATCCCTGTCTGACCAAGTGCAGAGAGAGCTGCTGACGACCAAGGTACCCAGAA 2657
QY 78 Leu----- 78
Db 2658 CT-GACGTAGTGCACACCGAGCTCCCGCGCGCGCGCGCGCGCGCGCGCTGA 2716
QY 78 ----- 78
Db 2717 CCCTCTGGGAGACCGTGTGTCTGAGACCTCAGGCTCACCGCGCGGTTTCCTCTG 2776
QY 78 ----- 78
Db 2777 TCTTGTGCCCAACTCTTGGGGGTCTGGGTCTGTCTTTCTTTTCTCTCTCTT 2836
QY 78 ----- 78
Db 2837 TTGGGGGGAACATTTTCTTTTCTTTCTTTCTTTGACTTCATGCTTGTCTTTCC 2896
QY 78 ----- 78

```


Db	2837	TTGGGGGAAAAAATTTTCTTTTCTTTTCTTTGACTTGGACTCATGCTCTGTTCTTCC	28936
QY	78	-----	78
Db	2897	ATCTTGAGCTTCCTGCCTTGCCTGTCTCTGGTGCAGCTTGGCGTCCCTGCTGTCTCG	29566
QY	78	-----	78
Db	2957	AATCTTG6CACGTCCTGGCCATCCGACGTACAGGACCCCTCTTCTCCCTCACGCCC	30166
QY	79	-----ArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaThr	92
Db	3017	CCCGCCCTCTCTGGCCCGAGGAGGTGATGAGAGAGATGATGAAGAGGTGAAGCCCTAC	30766
QY	93	LysSerGluLeuGluGluGlnLeuThrProValAlaGluGluThrArgAlaArgLysSer	112
Db	3077	CGCGAGGAGCTGGAGGGCGAGCTGGGGCCCTGACCCAGAGAGACCAAGCGGCTGTCC	31366
QY	113	LysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAspValCysGlyArg	132
Db	3137	AAGGAGCTGCAGGGCGGGCGAGCCCGGCTGGCGCCGAGATGAGAGACTGGCCAACTGC	31966
QY	133	LeuValGlnThrArgGlyGlyValAlaGlnAlaMetLeuAspGlnSerThrGluLeuArg	152
Db	3197	TTGGGTGCTCTACCCGACAGGAGTGACAAACATTTTGGCGACACCCAGAGAGCTGGCG	32566
QY	153	ValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAsp	172
Db	3257	AGCGCCCTGGCTTCCACACGCGCCAGCTGGCCAAACGGCTCTCTCGGACACCGAGGAC	33166
QY	173	LeuGlnLysCysLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeu	192
Db	3317	CTGCAGAGGCCCTGGCGCTGTCTACAGGGGGGCTGGCGAGGGCGCGAGCGCAGCTG	33766
QY	193	SerAlaIleArgGluArgLeuGlnGlyProLeuValGluGlnGlyArgValArgAlaIleThr	212
Db	3377	AGCGCCCTCGCCGAGGCCCTCGGGCCCTGTGTGAGACAGGGCGCATTCGGCCGCCAC	34366
QY	213	ValGlySerLeuAlaGlyGlnProLeuGlnGluArgAlaGlnAlaTyrGluArgLeu	232
Db	3437	CTGAGTACAGGCGCGGCGACGCGCTGCGGACGCGCGGAGACCTGGGGCCAGAACTG	34966
QY	233	ArgAlaArgMetGluGluMetGlySerArgThrArgAspArgLeuAspGluValLysGlu	252
Db	3497	CGCGAGCGGCTGGAGGAGATGGGACGCGGACCGCGGACCGCTGTGATGAGATGGCTAG	35566
QY	253	GlnValAlaGluValAlaArgAlaTyrLeuGlnGluGlnAlaGlnIleLeuGlnGlnAla	272
Db	3557	CAGCTGAGGAGAGGTGCCACCAAGTGGAGAGACGACGACGACCATTCGCTCGAGGCC	36166
QY	273	GluAlaPheGlnAlaArgLeuLysSerTyrPheGluProLeuValGluAspMetGlnArg	292
Db	3617	GAGGAGATTCCACGCGCTTCCAAAGGCTGGTTCGACCTCTGTGGAACATATAGCGCGC	36766
QY	293	GlnThrPheArgLeuValGluLysValAlaGlnAlaAlaValGly---ThrSerAlaAlaPro	311
Db	3677	CAGTGGCGGGGCTGGTGGAGAGATGACAGTGGCGGTGAGCATAGCTCTCCACCTCT	37366
QY	312	ValProSerAspAsn	316
Db	3737	GCGCCCACTGATTAAT	3751
RESULT 6			
US-08-726-306A-28			
: Sequence 28, Application US/08726306A			
: Patent No. 5958684			
: GENERAL INFORMATION:			
: APPLICANT: van Leeuwen, Frederik Willem			
: APPLICANT: Burbach, Johannes Peter Henri			
: APPLICANT: Grosveld, Franklin G.			
: TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS			
: NUMBER OF SEQUENCES: 189			
: CORRESPONDENCE ADDRESS:			

	ADDRESSSEE: Banner & Wilcoff, Ltd.	
	STREET: 1 Financial Center	
	CITY: Boston	
	STATE: MA	
	COUNTRY: US	
	ZIP: 02111	
	COMPUTER READABLE FORM:	
	MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage	
	OPERATING SYSTEM: PC-DOS/MS-DOS	
	SOFTWARE: WordPerfect 6.1	
	CURRENT APPLICATION DATA:	
	APPLICATION NUMBER: US/08/726,306A	
	FILING DATE: 02-Oct-1996	
	PRIOR APPLICATION DATA:	
	APPLICATION NUMBER: GB 95/20080.4	
	FILING DATE: 02-Oct-1995	
	PRIOR APPLICATION DATA:	
	APPLICATION NUMBER: US 60/009,832	
	FILING DATE: 01-Jan-1996	
	ATTORNEY/AGENT INFORMATION:	
	NAME: Williams, Ph.D., Kathleen M.	
	REGISTRATION NUMBER: 34,380	
	REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)	
	TELECOMMUNICATION INFORMATION:	
	TELEPHONE: (617) 345-9100	
	TELEFAX: (617) 345-9111	
	INFORMATION FOR SEQ ID NO: 28:	
	SEQUENCE CHARACTERISTICS:	
	LENGTH: 660 base pairs	
	TYPED: nucleic acid	
	STRANDEDNESS: double	
	TOPOLOGY: linear	
	MOLECULE TYPE: cDNA	
	US-08-726-306A-28	
	Alignment Scores:	
	Pred. No.: 3,23e-86 Length: 660	
	Score: 969.00 Matches: 195	
	Percent Similarity: 97.99% Conservative: 0	
	Best Local Similarity: 97.99% Mismatches: 4	
	Query Match: 60.98% Indels: 0	
	DB: 2 gaps: 0	
US-09-827-854-17 (1-317) x US-08-726-306A-28 (1-660)		
OY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal		20
Dd 62 ATGAAGGTTCCTGTGGCGTCCGTGCTGCATTCCTCGCAGAGATGCCAGGCCAAAGTG		122
OY 21 GluGlnAlaValAlaGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer		40
Dd 122 GAGCAAGCGGTGGAGACAGACC CGGAGCCGAGCTGCGCCAGCAGACCGAGTGGCAAGC		181
OY 41 GlyGlnAspGTPGluLeuAlaLeuGlyIleValPheTrpAspIleuAspGTPValGlnThr		60
Dd 182 GCCCGCCCTGGGAACCTGGCACCTGGCTTTGGGATTACTGGCCCTGGGTGCAGACA		241
OY 61 LeuSerGluGlnValAlaGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla		80
Dd 242 CTGTGTGACGAGGTGCAGAGAGAGCTGCTCAGCTCCCAAGTCAACCCAGAAGCATGAGGCG		301
OY 81 LeuMetAspGluThrMetLysGluLeuLeuValIleThrLysSerGluLeuGlnAlaGlnAla		100
Dd 302 CTGATGAGCAGAACCATTAAGAGTTGAAGCCCTTCAAAATCGGAACATGAGGAACAACATG		361
OY 101 ThrProValAlaGluGluThrArgAlaLeuSerLysGluLeuGlnAlaGlnAla		120
Dd 362 ACCCGGTAGCGGAGAGAACCGGGCAGCGCTGTCCAAAGGAGCTGACGAGCGCGCAGGCC		421
OY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyIleValGlnIleThrArgGlyGluVal		140
Dd 422 CGCGTGGCGCGGACATGAGAGAGCTGTGGCGCGCTGTGTGCATGACCGCGGAGAGTG		481

MOLECULE TYPE: CDNA
US-09-431-613-24

Alignment Scores:

Pred. No.:	5,73e-28	Length:	252
Score:	367.00	Matches:	81
Percent Similarity:	96.43%	Conservative:	0
Best Local Similarity:	96.43%	Mismatches:	3
Query Match:	23.10%	Indels:	1
DB:	4	Gaps:	0

US-09-827-854-17 (1-317) x US-09-431-613-24 (1-252)

QY 109 AlaArgLeuSerIysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
|||||
DB 2 GCACGGCTGTCCCAAGAGCTGCAGCGCGCCGACGCCGCTGGCGCGGACATGAGAGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuAspGlnSerThr 148
|||||
DB 62 GTGTGGCGC-CGCTGTGTGACGTACCGCGGAGGTGACAGCCATGCTCGGCGCACAGACACC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
|||||
DB 121 GAGGACCTGGGGTGGCGCTCGCTCCACCTCGCAGCTGCGTAAGCGGCTCTCCGCC 180

QY 169 AspAlaAspAspLeuGlnLysCysLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
|||||
DB 181 GATGCCGATGACCTGCAGAACTCCCTGGCAGTGTACACAGCCGGCGCGCGAGGCGGCC 240

QY 189 GluArgGlyLeu 192
|||||
DB 241 GAGCGCGGCGCTC 252

RESULT 11
US-09-504-245-24
Sequence 24, Application US/09504245
Patent No. 6221605
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe LLP
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/504,245
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002J
TELECOMMUNICATION INFORMATION:
TELEPHONE: 858-450-8400
TELEFAX: 858-567-5360
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-504-245-24

Alignment Scores:

Pred. No.:	5,73e-28	Length:	252
Score:	367.00	Matches:	81
Percent Similarity:	96.43%	Conservative:	0
Best Local Similarity:	96.43%	Mismatches:	3
Query Match:	23.10%	Indels:	1
DB:	4	Gaps:	0

US-09-827-854-17 (1-317) x US-09-504-245-24 (1-252)

QY 109 AlaArgLeuSerIysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
|||||
DB 2 GCACGGCTGTCCCAAGAGCTGCAGCGCGCCGACGCCGCTGGCGCGGACATGAGAGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuAspGlnSerThr 148
|||||
DB 62 GTGTGGCGC-CGCTGTGTGACGTACCGCGGAGGTGACAGCCATGCTCGGCGCACAGACACC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
|||||
DB 121 GAGGACCTGGGGTGGCGCTCGCTCCACCTCGCAGCTGCGTAAGCGGCTCTCCGCC 180

QY 169 AspAlaAspAspLeuGlnLysCysLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
|||||
DB 181 GATGCCGATGACCTGCAGAACTCCCTGGCAGTGTACACAGCCGGCGCGCGAGGCGGCC 240

QY 189 GluArgGlyLeu 192
|||||
DB 241 GAGCGCGGCGCTC 252

RESULT 12
US-09-287-682-24
Sequence 24, Application US/09287682
Patent No. 6235478
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002E
TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-682-24

Alignment Scores:
Pred. No.: 5,73e-28
Score: 367.00
Percent Similarity: 96.43%
Best Local Similarity: 96.43%
Query Match: 23.10%
Matches: 81
Conservative: 0
Mismatches: 3
Indels: 1
Gaps: 0

US-09-827-854-17 (1-317) x US-09-287-682-24 (1-252)

QY 109 ALAARGLEUSERYSGIULEUGINALAALAGLNAALARGLEUGLYLAASPMETGLUASP 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGGCGCGCCGCTGGCGCGGACATGGAGGAC 61

QY 129 VALCYSGLYARGLEUVALGINTYRARGLYGLUVALGLNALMETLEUASPGINSETRH 148
DB 62 GTGTGGCCGCGCTGTGAGTACCGCGGAGGTGCAAGCCATGCTGGCCAGAGCACC 120

QY 149 GLUGLUENURGYVALARGLEUALASERHISLEUARGLYSLEUARGLYSARGLEUARG 168
DB 121 GAGGAGCTGGGGGTGGCCCTGCCTCCACTGCGCAAGCTGCTAAGCGGCTCTCCGC 180

QY 169 ASPALASPAPLEUGLNLSCYSLLEUALAVALTYRGINALAGLYLAARGGLUGLYALA 188
DB 181 GATGCCGATGACCTGCAGAGAGTCCCTGCGAGTACACAGCGCGGCGCGCGAGGCGCC 240

QY 189 GLUARGGLYLEU 192
DB 241 GAGCGGGGCTC 252

RESULT 13
US-09-287-679-24
Sequence 24, Application US/09287679
Patent No. 6258538
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287, 679
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-679-24

Alignment Scores:
Pred. No.: 5,73e-28
Score: 367.00
Percent Similarity: 96.43%
Best Local Similarity: 96.43%
Query Match: 23.10%
Matches: 81
Conservative: 0
Mismatches: 3
Indels: 1
Gaps: 0

US-09-827-854-17 (1-317) x US-09-287-679-24 (1-252)

QY 109 ALAARGLEUSERYSGIULEUGINALAALAGLNAALARGLEUGLYLAASPMETGLUASP 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGGCGCGCCGCTGGCGCGGACATGGAGGAC 61

QY 129 VALCYSGLYARGLEUVALGINTYRARGLYGLUVALGLNALMETLEUASPGINSETRH 148
DB 62 GTGTGGCCGCGCTGTGAGTACCGCGGAGGTGCAAGCCATGCTGGCCAGAGCACC 120

QY 149 GLUGLUENURGYVALARGLEUALASERHISLEUARGLYSLEUARGLYSARGLEUARG 168
DB 121 GAGGAGCTGGGGGTGGCCCTGCCTCCACTGCGCAAGCTGCTAAGCGGCTCTCCGC 180

QY 169 ASPALASPAPLEUGLNLSCYSLLEUALAVALTYRGINALAGLYLAARGGLUGLYALA 188
DB 181 GATGCCGATGACCTGCAGAGAGTCCCTGCGAGTACACAGCGCGGCGCGCGAGGCGCC 240

QY 189 GLUARGGLYLEU 192
DB 241 GAGCGGGGCTC 252

RESULT 14
US-09-397-766-24
Sequence 24, Application US/09397766
Patent No. 6268144
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/397,766
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-20021
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-397-766-24

Alignment Scores:
Pred. No.: 5,73e-28 Length: 252
Score: 367.00 Matches: 81
Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 23.10% Indels: 1
DB: Gaps: 0

US-09-827-854-17 (1-317) x US-09-397-766-24 (1-252)

QY 109 AlaAargLeuSerLyGluLeuGlnAlaAargLeuGlyAlaAspMetGluasp 128
|||||
DB 2 GCACGCGCTGTCCAGACGCTGCAGGCGCGCGCGCTGGCGCGGACATGGAGGAC 61
QY 129 ValCysGlyAargLeuValGlnTyrAargGlyGluValGlnAlaMetLeuAspGlnSerThr 148
|||||
DB 62 GGTGTGCG-CGCGTGTGTGAGTACCGCGCGGAGGTGCAGGCGCATGCTCGGCCAGACACC 120
QY 149 GluGluLeuAargValAargLeuAlaSerHisLeuAargLysLeuAargLysAargLeuAarg 168
|||||
DB 121 GAGGAGCTGCGGGTGGCTCGCTCCACCTGCGCAAGCTCGTAAGCGGCTCTCCGCG 180
QY 169 AspAlaAspAspLeuGlnLysCysLeuAlaValTyrGlnAlaGlyAlaAargGluGlyAla 188
|||||
DB 181 GATGCCGATGACCTGCAGAACTCTGGCAGTGTACAGCGCGGCGCGCGGAGGCGCC 240
QY 189 GluAargGlyLeu 192
|||||
DB 241 GAGCGGCGGCTC 252

RESULT 15

US-09-287-681-24
Sequence 24, Application US/09287681
Patent No. 627573

GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,681
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-287-681-24

Alignment Scores:
Pred. No.: 5,73e-28 Length: 252
Score: 367.00 Matches: 81
Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 23.10% Indels: 1
DB: Gaps: 0

US-09-827-854-17 (1-317) x US-09-287-681-24 (1-252)

QY 109 AlaAargLeuSerLyGluLeuGlnAlaAargLeuGlyAlaAspMetGluasp 128
|||||
DB 2 GCACGCGCTGTCCAGACGCTGCAGGCGCGCGCGCTGGCGCGGACATGGAGGAC 61
QY 129 ValCysGlyAargLeuValGlnTyrAargGlyGluValGlnAlaMetLeuAspGlnSerThr 148
|||||
DB 62 GGTGTGCG-CGCGTGTGTGAGTACCGCGCGGAGGTGCAGGCGCATGCTCGGCCAGACACC 120
QY 149 GluGluLeuAargValAargLeuAlaSerHisLeuAargLysLeuAargLysAargLeuAarg 168
|||||
DB 121 GAGGAGCTGCGGGTGGCTCGCTCCACCTGCGCAAGCTCGTAAGCGGCTCTCCGCG 180
QY 169 AspAlaAspAspLeuGlnLysCysLeuAlaValTyrGlnAlaGlyAlaAargGluGlyAla 188
|||||
DB 181 GATGCCGATGACCTGCAGAACTCTGGCAGTGTACAGCGCGGCGCGGAGGCGCC 240
QY 189 GluAargGlyLeu 192
|||||
DB 241 GAGCGGCGGCTC 252

Search completed: March 14, 2003, 20:18:46
Job time : 40.7343 secs


```

Db      661 CCCCTGTTGTAACAGGCGCCGCTGCGGCGCCACTGTGGCTCCCTGGCCGCGCAAGCCG 720
Qy      221 LeuGluGluArgAlaGlnAlaIleThrPgluArgLeuArgAlaIleThrPgluGluGly 240
Db      721 CTACAGAGACGGGCGCCAGGCTGTGGCGAGCGGCTGTGGCGGATGTGAGAGATGGGC 780
Qy      241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db      781 AGCCGAGACCCGCGACCGCTGAGCAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Qy      261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db      841 CTGGAGGAGACAGCCCGACAGATAGCCCTGCGAGCGCGAGCCCTTCCAGGCGCCCTCAAG 900
Qy      281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnIleThrAlaGlyLeuValGlyLys 300
Db      901 ACCTGTTTGGAGCCCTGTGTGAGACATGACAGCGCATGTGGCGGCGGTGTGTGAGAGAG 960
Qy      301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db      961 GTGCAAGCTGCGGTGGGCGACAGCGCCGCTGTGTGCCAGGACATATAC 1011

RESULT 3
; US-09-870-759-129
; Sequence 129, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 129
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(1014)
; OTHER INFORMATION:
US-09-870-759-129

Alignment Scores:
Pred. No.: 1,05e-135
Score: 1570.00
Percent Similarity: 99.37%
Best Local Similarity: 99.37%
Query Match: 98.80%
DB: 9 Gaps: 0

US-09-827-854-17 (1-317) x US-09-870-759-129 (1-1156)
Qy      1 MetLysValLeuThrAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db      61 ATGAAGGTTCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGT 120
Qy      21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGluThrGluThrPgluIns 40
Db      121 GAGCAAGCGGTGAGACAGAGCGGAGCGGAGCTGCGGACAGACAGACAGAGCGAGAGC 180
Qy      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
Db      181 GCCCAGCGCTGGAGACTGCGACTGCTGCTTTGGGATTACTGCGCTGCGTGCAGAGACA 240
Qy      61 LeuSerGlnGlnValGlnGluGluLeuSerGlnValThrGlnGluLeuArgAla 80
Db      241 CTGTCTGACAGAGTGAAGAGAGCTGCTCAGCTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 300
Qy      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu 100

```

```

Db      301 CTGATGAGACAGACCATGAAGAGTTGAAGGCGCTCAAAATCGAGACTGGAGAGAACACTG 360
Qy      101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db      361 ACCCGGTGGCGAGAGAGCGGGGACAGCTGTCTCCAAAGACTCCAGCGCGGAGAGCC 420
Qy      121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGluTyrArgGlyGluVal 140
Db      421 CGGCTGTGGCGCGGACATGAGAGACGTGTGCGGCGCTGTGTGCTACGTACCGCGGAGGTG 480
Qy      141 GlnAlaMetLeuAspGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db      481 CAGCGCATCTCGGCGCAGAGACCGAGAGCTGCGGGTGTGCGCTGCGCTCCACCTGCGC 540
Qy      161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db      541 AAGCTGCGTAAACGGGCTCCCGCGATGCGATGACCTGCAGAAAGCGCTGCGCAGTGTAC 600
Qy      181 GlnAlaGlyAlaArgGluGluValAlaGluArgGlyLeuSerAlaIleArgLeuArgLeuGly 200
Db      601 CAGCGCGGGGCGCGGAGGCGCGGCGGCTCTAGCGCATCCGCGAGCGCGCTGGGG 660
Qy      201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db      661 CCCCTGTGTGAACAGAGCGCGGCTGTGGCGGCGGCGGCTGTGTGTGTGTGTGTGTGTGTGT 720
Qy      221 LeuGlnGluArgAlaGlnAlaThrPgluArgLeuArgAlaArgMetGluGlyGly 240
Db      721 CTACAGAGACGGGCGCGAGGCTGTGGGCGAGCGGCTGTGGCGGCGGATGAGAGATGGGC 780
Qy      241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db      781 ACCCGAGACCGGAGCGGCTGTGAGCAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Qy      261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db      841 CTGGAGAGACAGCCCGACAGATAGCCCTGCGAGCGCGAGCGCTTCCAGGCGCGGCTCAAG 900
Qy      281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnIleThrAlaGlyLeuValGlyLys 300
Db      901 ACCTGTTTGGAGCCCTGTGTGAGACATGACAGCGCATGTGGCGGCGGTGTGTGAGAGAG 960
Qy      301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db      961 GTGCAAGCTGCGGTGGGCGACAGCGCCGCTGTGTGCCAGGACATATAC 1011

RESULT 4
US-09-802-640-17
; Sequence 17, Application US/09802640
; Publication No. US20030036057A1
; GENERAL INFORMATION:
; APPLICANT: Braun, Andreas
; APPLICANT: Bonsal Aruna
; APPLICANT: Kleyen Patrick
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: 24736-2048
; CURRENT APPLICATION NUMBER: US/09/802,640
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(1014)
; OTHER INFORMATION: Nucleotide Sequence encoding apolipoprotein E
; OTHER INFORMATION: (APOE)
US-09-802-640-17

```


[illegible]

Db 366 CTTGTACAGAGGTGTGAGAGGAGCTGTACAGCTCCAGAGTACCCAGAGAACTGAGGGCG 425
 QY 81 LeuMeVAspGluThrMetLysGluLeuLysAlaThrLysSerGluLeuGluGluGluLeu 100
 Db 426 CTTGATGGAGACACCTTGAAGAGTTGAAGAGCCTACAAATCGGAAGTGGAGAACTG 485
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGluAlaAlaGluAla 120
 Db 486 ACCCGCGTGGCGAGAGAGACCGGGCACGGCTGTCCAGAGAGCTGCAGGGCGGCGAGGCC 545
 QY 121 ArgLeuGlyAlaAspMetGluAspValLysGlyArgLeuValGluThrArgGlyGluVal 140
 Db 546 CGGCTGGGGCGGACACTGGAGGACGTGTGCGGGCGGCTGTGTGACGAAACCCGGCGAGGTG 605
 QY 141 GlnAlaMetLeuAspGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 606 CAGGCCATCTCTGCGCCACAGACCCAGAGAGCTGTGGGTGGCGCTCGCTTCCACCTGGCG 665
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGluLysCysLeuAlaValThr 180
 Db 666 AAGCTGCGTAAAGCGGCTCTCCGCGATGCGGATGACCTGTGAGAAAGCGCTGGCAGGTGTAC 725
 QY 181 GlnAlaGlyAlaArgGluGluAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db 726 CAGGCGGGGGCCCCGAGGGCGCCGAGCGGCGCTCAGGCGCCATCCGCGAGCCCTGGGG 785
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 786 CCCCTGTGTGAACAGAGGGCGCGCTGCGGGCGCCACCTGTGGTCTCCCTGGCGGCGACCG 845
 QY 221 LeuGluGluArgAlaGlnAlaThrPglGlyArgArgLeuArgAlaArgMetGluMetGly 240
 Db 846 CTACAGAGACGGGCCCCAGGCGCTGGGGCGAGGGCGCTGGCGCGGATGGAGAGATGGGC 905
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 Db 906 ACCCGGACCCCGGACCGCTGTGACGAGGTGAAGAGACAGGTGGCGAGGCGCGCCAG 965
 QY 261 LeuGluGluAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 Db 966 CTGGAGAGACAGGCCAGCAGATACGCTGGACGGCCGAGGCGCTTCCAGAGCGCCCTCAAG 1025
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 Db 1026 ACGTGGTTGACACCCCTGTGTGAAGACATGACAGCGCCAGTGGGGCGGGCGGTGGAGAG 1085
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 1086 GTGACAGGCTGCGGTGGCACACAGCGCGCCCTGTGCCACAGCAGCAATCATC 1136
 RESULT 7
 : US-09-827-854-12
 : Sequence 12, Application US/09827854
 : Patent No. US20020123093A1
 : GENERAL INFORMATION:
 : APPLICANT: Kypros, Kyriakos E.
 : APPLICANT: Zannis, Vassilis
 : TITLE OF INVENTION: Compounds and methods for lowering
 : FILE REFERENCE: 07180/004003
 : CURRENT APPLICATION NUMBER: US/09/827,854
 : CURRENT FILING DATE: 2001-04-05
 : PRIOR APPLICATION NUMBER: US 09/679,088
 : PRIOR FILING DATE: 2000-10-04
 : PRIOR APPLICATION NUMBER: US 09/544,386
 : PRIOR FILING DATE: 2000-04-06
 : NUMBER OF SEQ ID NOS: 19
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 12
 : LENGTH: 1156
 : TYPE: DNA
 : ORGANISM: Homo sapiens


```

; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2244
; LENGTH: 1157
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M12529
US-09-880-107-2244
```

Alignment Scores:

Pred. No.:	7.03e-135	Length:	1157
Score:	1561.00	Matches:	313
Percent Similarity:	98.74%	Conservative:	0
Best Local Similarity:	98.74%	Mismatches:	4
Query Match:	98.24%	Indels:	0
DB:	10	Gaps:	0

US-09-827-854-17 (1-317) x US-09-880-107-2244 (1-1157)

```

QY 1 MetlyValleuTrpaAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||||||
DB 62 ATGAGGTTCTGTGGGCTGGTGTCTGTGCATCTCTGGCAGAGATCCAGGCAAGTGT 121
    |||||||
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
    |||||||
DB 122 GAGCAACGGGTGGAGACGAGCCGAGCCGAGCTCGCCAGCAGACCGAGTGGCAGACG 181
    |||||||
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
    |||||||
DB 182 GGCAGGCGCTGGAGACTGGACTGGCTGCTTTGGATTACCTGGCTGGGTGCAGACA 241
    |||||||
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
    |||||||
DB 242 CTGCTGAGCAGAGTGCAGAGAGAGCTGCTCAGCTCCCAATCCCAAGAACTGAGGGCG 301
    |||||||
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
    |||||||
DB 302 CTGATGGACGACGACCATGAAGAGTTGTAAGCCTACAAATCGGAACCTGGAGAACTG 361
    |||||||
QY 101 ThrProValGluGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
    |||||||
DB 362 ACCCGGTTAGCGAGAGAGACCGGGCCGCTGTCCAAAGAGCTGCCAGACGGCGCAGGCC 421
    |||||||
QY 121 ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
    |||||||
DB 422 CGGCTGGCGCCGACATGAGAGAGCTGTGGCGCCGCTGTGTCAGTACCGCGGAGAGTGT 481
    |||||||
QY 141 GlnAlaMetLeuAspGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
    |||||||
DB 482 CAGGCGCATGCTCGCGCACCGACCGAGAGCTGGGGTGGCTCGCTCCGCTCCACCTGCGC 541
    |||||||
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
    |||||||
DB 542 AACCTGCGTAAAGCGGCTCTCCCGGATCCGATGACCTCGAGAAAGCCCTGGGCGAGTAC 601
    |||||||
QY 181 GlnAlaGlyAlaArgGluGlnGluAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
    |||||||
DB 602 CAGGCGGCGGCGCGGAGCGCGGAGCGGCGGCTCAGGCCCATCCGCGAGCGGCTGGGG 661
    |||||||
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
    |||||||
```

```

DB 662 CCCCTGTGTGAACAGGCGCCCGGCGCCGACATCTGTGGCTTCCTGGCGCCGACCGC 721
    |||||||
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyLysArgLeuArgAlaArgMetGluMetGly 240
    |||||||
DB 722 CTACAGAGCGGGGCCAGGCTGTGGGCGAGCGGCTGGCGCGGATGAGAGATGGCG 781
    |||||||
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
    |||||||
DB 782 AGTGGAGACCCGCGACCGCTCGTGAAGAGGTGAGAGAGAGTGGCGGAGGTGGCGCCAG 841
    |||||||
QY 261 LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
    |||||||
DB 842 CTGGAGAGCAGGCGCCGACGAGATACGCTCAGCGCGAGGCTTCCAGGCGCCCTCAG 901
    |||||||
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
    |||||||
DB 902 AGCTGTTTCAGACCCCTCGTGTGAAGACATCGACGCCAGTGGCGGCTGTGTGAGAG 961
    |||||||
QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
    |||||||
DB 962 GTGAGGCTCGCTGGGCGACAGCGCGCCCTGTGGCCAGCAGCAATCAC 1012
    |||||||
```

RESULT 11

```

US-09-827-854-7
; Sequence 7, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassilis
; APPLICANT: Kyriacos, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-7
```

Alignment Scores:

Pred. No.:	1.32e-134	Length:	1156
Score:	1558.00 <td>Matches:</td> <td>314</td>	Matches:	314
Percent Similarity:	99.05% <td>Conservative:</td> <td>0</td>	Conservative:	0
Best Local Similarity:	99.05% <td>Mismatches:</td> <td>3</td>	Mismatches:	3
Query Match:	98.05% <td>Indels:</td> <td>0</td>	Indels:	0
DB:	10	Gaps:	0

US-09-827-854-17 (1-317) x US-09-827-854-7 (1-1156)

```

QY 1 MetlyValleuTrpaAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||||||
DB 61 ATGAGGTTCTGTGGGCTGGTGTCTGTGCATCTCTGGCAGAGATCCAGGCAAGTGT 120
    |||||||
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
    |||||||
DB 121 GAGCAACGGGTGGAGACGAGCGCGGAGCCGACCTGCGCCAGCAGACCGAGTGGCAGAGC 180
    |||||||
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
    |||||||
DB 181 GGCAGGCGCTGGAGACTGTGGCTGCTTTGGATTACCTGGCTGGGTGCAGACAGA 240
    |||||||
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
    |||||||
DB 241 CTGCTGAGCAGAGTGCAGAGAGAGCTGCTCAGCTCCAGGTACCCAGGAACCTGAGGGCG 300
    |||||||
```


Db 396 ATGAGGTTCTGTGGGCTGGCTGCTGTCACATTCTCGCAGAGATCCAGGCCAAGCTG 337
QY 21 GtuglnAlaValGluThrGluProGluProGluLeuArgIngluThrGluTrpGlnSer 40
|||||
Db 336 GAGCAACCGGTGGAGACAGAGCCGAGCCNCACTGGCGCAGACAGCCAGTGCGCAGAGC 277
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
Db 276 GGGCAGCGCTGGAGACTGGACGTGGCTTGGATTACCTGGCTGGCTGGCGAGACA 217
QY 61 LeuSerGluGlnValGlnGluLeuLeu-SerSerGlnValThrGlnGluLeuArgAl 80
|||||
Db 216 CTCTGTGAGAGGTGCAGAGGAGCTGCTGACCTCCAGCTCCAGCAGGAACTGAGGGC 157
QY 80 AleuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLe 100
|||||
Db 156 GCTGATGAGACAGACCATGAGAGGTTGAGAGGCTTCAAAATCGGAATGAGAGAACT 97
QY 100 uThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAl 120
|||||
Db 96 GACCCGCGTGGCGGAGAGACGCGGCGACGCTGTCCAAAGAGCTGCAGCGCGCGAGN 37
QY 120 aArgLeuGlyAlaAspMetGluAspVal 129
|||||
Db 36 CCGGCTGGCGCGACATGAGGAGCGTG 9
RESULT 15
US-09-880-107-2491/c
; Sequence 2491, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2491
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 N33009
; NAME/KEY: unsure
; LOCATION: (1)-(478)
; OTHER INFORMATION: n = a o r c o r g o r t
US-09-880-107-2491
Alignment Scores:
Pred. No.: 2,16e-49 Length: 478
Score: 625.00 Matches: 126
Percent Similarity: 97.698 Conservative: 1
Best Local Similarity: 96.928 Mismatches: 2
Query Match: 39.338 Indels: 1
DB: 10 Gaps: 0
US-09-827-854-17 (1-317) x US-09-880-107-2491 (1-478)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
Db 396 ATGAGGTTCTGTGGGCTGGCTGCTGTCACATTCTCGCAGAGATCCAGGCCAAGCTG 337
QY 21 GtuglnAlaValGluThrGluProGluProGluLeuArgIngluThrGluTrpGlnSer 40
|||||
Db 336 GAGCAACCGGTGGAGACAGAGCCGAGCCNCACTGGCGCAGACAGCCAGTGCGCAGAGC 277
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
Db 276 GGGCAGCGCTGGAGACTGGACGTGGCTTGGATTACCTGGCTGGCTGGCGAGACA 217
QY 61 LeuSerGluGlnValGlnGluLeuLeu-SerSerGlnValThrGlnGluLeuArgAl 80
|||||
Db 216 CTCTGTGAGAGGTGCAGAGGAGCTGCTGACCTCCAGCTCCAGCAGGAACTGAGGGC 157
QY 80 AleuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLe 100
|||||
Db 156 GCTGATGAGACAGACCATGAGAGGTTGAGAGGCTTCAAAATCGGAATGAGAGAACT 97
QY 100 uThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAl 120
|||||
Db 96 GACCCGCGTGGCGGAGAGACGCGGCGACGCTGTCCAAAGAGCTGCAGCGCGCGAGN 37
QY 120 aArgLeuGlyAlaAspMetGluAspVal 129
|||||
Db 36 CCGGCTGGCGCGACATGAGGAGCGTG 9

Search completed: March 15, 2003, 03:04:37
Job time : 82.8753 secs

QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
Db 276 GGGCAGCGCTGGAGACTGGACGTGGCTTGGATTACCTGGCTGGCTGGCGAGACA 217
QY 61 LeuSerGluGlnValGlnGluLeuLeu-SerSerGlnValThrGlnGluLeuArgAl 80
|||||
Db 216 CTCTGTGAGAGGTGCAGAGGAGCTGCTGACCTCCAGCTCCAGCAGGAACTGAGGGC 157
QY 80 AleuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLe 100
|||||
Db 156 GCTGATGAGACAGACCATGAGAGGTTGAGAGGCTTCAAAATCGGAATGAGAGAACT 97
QY 100 uThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAl 120
|||||
Db 96 GACCCGCGTGGCGGAGAGACGCGGCGACGCTGTCCAAAGAGCTGCAGCGCGCGAGN 37
QY 120 aArgLeuGlyAlaAspMetGluAspVal 129
|||||
Db 36 CCGGCTGGCGCGACATGAGGAGCGTG 9

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:09:47 ; Search time 1289.69 Seconds

(without alignments)
3980.771 Million cell updates/sec

Title: US-09-827-854-17
Perfect score: 1589
Sequence: 1 MKVLMAALVFLAGCOAKV.....VEKYQAAVGTSAAPVSDNH 317

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus.p2n.model -DEV=xlp
-O=/cgn2.1/USP10.spool/US09827854/runat.11032003.101610.27495/app.query.fasta.1.3576
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-ONITS-bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-OCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=5001 -MINLEN=0 -MAXLEN=2000000000
-USER=US09827854.ecgn.1.1.8826_@runat.11032003.101610.27495 -NCPU=6 -ICPU=3
-NO_XIPXY -NO_MMAP -LARGEQUERY -NEGSCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :
EST : *
1: em_estba : *
2: em_estbam : *
3: em_estlin : *
4: em_estlmu : *
5: em_estlov : *
6: em_estlpl : *
7: em_estlro : *
8: em_hlc : *
9: gb_estl1 : *
10: gb_estl2 : *
11: gb_hlc : *
12: gb_est3 : *
13: gb_est4 : *
14: gb_est5 : *
15: em_estfun : *
16: em_estlom : *
17: gb_gss : *
18: em_gss_hum : *
19: em_gss_inv : *
20: em_gss_dln : *
21: em_gss_vrt : *
22: em_gss_fun : *
23: em_gss_mam : *
24: em_gss_mus : *
25: em_gss_other : *
26: em_gss_pro : *
27: em_gss_rtd : *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1275.5	80.3	1027	13	B1670367	B1670367 603292738
2	1239	78.0	842	12	B6763371	B6763371 602735433
3	1233	77.6	933	13	B1668318	B1668318 603295681
4	1205	75.8	938	12	B6761746	B6761746 602717942
5	1202.5	75.7	942	13	B1600906	B1600906 603249241
6	1188	74.8	900	13	BM042094	BM042094 603615713
7	1187.5	74.7	927	12	B6472299	B6472299 602513830
8	1184.5	74.5	922	13	B1597743	B1597743 603248609
9	1175.5	74.0	817	12	B6774871	B6774871 602649975
10	1174	73.9	811	13	B1600563	B1600563 603244936
11	1160.5	73.0	845	12	B6829472	B6829472 602763768
12	1137	71.6	790	12	B6707147	B6707147 602670283
13	1134	71.4	919	13	B1551475	B1551475 603194314
14	1124	70.7	757	13	BM042228	BM042228 603616186
15	1114.5	70.1	907	12	B6706129	B6706129 602669093
16	1113	70.0	706	14	BM728696	BM728696 UI-E-EJ0-
17	1113	70.0	741	12	B6762924	B6762924 602735153
18	1111	69.9	1110	11	AK010261	AK010261 Mus. muscu
19	1109	69.8	804	12	B6702752	B6702752 602684616
20	1108	69.7	834	13	BM042676	BM042676 603616028
21	1105.5	69.6	808	13	B1668329	B1668329 603295692
22	1102	69.4	855	13	B161362	B161362 602865769
23	1099	69.2	803	13	B1670350	B1670350 603292721
24	1096	69.0	782	12	B6716776	B6716776 602678182
25	1096	69.0	812	13	B1601551	B1601551 603250760
26	1096	69.0	965	14	B6677266	B6677266 AGENCOURT
27	1093	68.8	797	12	B1603668	B1603668 603251982
28	1090.5	68.6	797	13	B6715366	B6715366 602675503
29	1090	68.6	846	13	B1159757	B1159757 602863509
30	1089	68.5	796	13	BM042153	BM042153 603615790
31	1084	68.2	757	12	B6707750	B6707750 602671277
32	1084	68.2	794	12	B1601279	B1601279 603245010
33	1084	68.2	954	12	BF967543	BF967543 602287404
34	1082	68.1	790	13	B1551066	B1551066 603196425
35	1081	68.0	790	13	B1551811	B1551811 603192716
36	1077.5	67.8	891	13	B6708414	B6708414 602672374
37	1076.5	67.7	891	13	B1548292	B1548292 603189930
38	1070.5	67.4	802	13	B1458355	B1458355 603198705
39	1069	67.3	1100	14	BM914382	BM914382 AGENCOURT
40	1068.5	67.2	914	13	B1603523	B1603523 603244424
41	1068	67.2	748	13	B1553085	B1553085 603196058
42	1067.5	67.2	930	12	BF967857	BF967857 6028287176
43	1063.5	66.9	812	12	B6769968	B6769968 602739654
44	1063	66.9	688	13	B1544886	B1544886 603242328
45	1062.5	66.9	803	12	B6709360	B6709360 602673385

ALIGNMENTS

RESULT 1
LOCUS B1670367 1027 bp mRNA EST 12-SEP-2001
DEFINITION 603292738F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5312024 5',
ACCESSION B1670367
VERSION B1670367.1 GI:15584600
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1027)
NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov

Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shluraki
Toshlyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

Plate: LLM11790 row: 1 column: 09
High quality sequence stop: 845.

FEATURES

Location/Qualifiers

1..1027
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5312024"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTVA-3',
size-selected for average insert size 2.3 kb and
normalized to R0T 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 194 a 308 c 400 g 125 t
ORIGIN

Alignment Scores:

Pred. No.: 1.09e-124 Length: 1027
Score: 1275.50 Matches: 290
Percent Similarity: 90.248 Conservative: 6
Best Local Similarity: 88.418 Mismatches: 21
Query Match: 80.27% Indels: 12
DB: 13 Gaps: 2

US-09-827-854-17 (1-317) x B1670367 (1-1027)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
DB 26 ATGAGAGTTCTGTGGCTGTGCTGTGCTGACATCTCTGCGCAGAGATGCCAGGCAAGTGT 85
QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnThrPglInser 40
|||||
DB 86 GAGCAAGCGGTGAGACAGAGCCGAGCCGAGCTGCGCAGACAGACCGAGTGGCAGAGC 145
QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
|||||
DB 146 GGGCAGCGGTGGGAGACTGGGACTGGCTGCTTTGGGATTACTGGCTGGTGGCGAGACA 205
QY 61 LeuSerGlnGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgGln 80
|||||
DB 206 CTGTCTGAGAGAGGTGACGAGGAGCTGCTCAGCTCCAGGTACCCAGGAACTGAGGGCG 265
QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaThrLysSerGlnLeuGlnGlnGln 100
|||||
DB 266 CTGATGAGAGAGACATGAGAGGAGTTGAAGGCTTACAAATCGGAATGGAGGAGCAACTG 325
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
|||||
DB 326 ACCCGGTTGGCGGAGAGACGCGGCGGCTCTCAAGAGCTGAGCGGCGGCGAGGCC 385
QY 121 ArgLeuGlyAlaAspMetGlnLysValCysGlyArgLeuValGlnThrArgGlyGlnVal 140
|||||
DB 386 CGGCTTGGCGGCGGACATGAGAGGACGTGTGCGCGCTGTTGACATGACCGCGGCGAGTGT 445
QY 141 GlnAlaMetLeuAspGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
|||||

DB 446 CAGGCCATGCTCGGCCAGAGACCCAGAGAGCTGGGGTGGCTCCGCTCCACCTGCGCC 505
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
|||||
DB 506 AAGCTGGTGAAGGGGCTCCCTCCGAGATGCCGATGACCTGACAGAGCGGCTGGCAGTAC 565
QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
|||||
DB 566 CAGCGCGGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 625
QY 200 YProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
|||||
DB 626 CCGCTGTGTGAGACAGGTCCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 665
QY 220 OleuGlnGlnArgAlaGlnAlaTrpGlnGlnArg-LeuArgAlaArgMetGlnGln-Met 239
|||||
DB 686 GCTACAGAGACGGGCGCAAGCGCTGGGGCGAGCGGCTTGCGCGCGGATGAGAGAGAG 745
QY 240 GlySerArg-ThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGln-ValArg 259
|||||
DB 746 GCGACGCGGAGCCCGGACCGCTGACCGAGGTGAAGAGACAGTGGCGGAGGTGCGCGC 805
QY 259 LysLeu---GlnGlnGlnAlaGlnGln-IleArgLeuGln-AlaGlnAlaPheGlnAl 277
|||||
DB 806 CCCAGCTGAGACGAGCCAGCCAGCAGATGAGCTGCGAGGCGCGAGGCTTCTAGGC 865
QY 277 AatGLeuLysSerTrpPheGlnPro-----LeuValGlnAspMetGlnArgGlnTrpAl 295
|||||
DB 866 CCGGCTCAGAGAGCTGGTGGTGAACCCCTGGCTGTGAACGACATGAGCGCGGCGGCAAGTGG 925
QY 295 agLysLeuValGlnLysVal-GlnAlaAlaValGlnThrSerAlaAlaProValProSer 315
|||||
DB 926 CGGCTGTGGAAGAGGTGCGGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 984
QY 315 spasnHis 317
DB 985 ACAATCAT 992

RESULT 2
BG763371 842 bp mRNA linear EST 15-MAY-2001
LOCUS 602735433F1 NIH_MGC_49 Homo sapiens CDNA clone IMAGE:4860585 5',
DEFINITION

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov

Tissue Procurement: ATCC/DC/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov
Plate: LLM1718 row: b column: 10
High quality sequence stop: 817.

FEATURES

Source

1..842
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4860585"
/clone_lib="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"

Db 195 GGCACGCGCTGGAGACTGGCAGCTGGTCTTTTGGATTACCTGCCCTGGGTGCAGACA 254
 QY LeuSerGluGlnValGlnGluGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 Db 255 CTGTCTGACGAGGTGCAGAGAGACTGCTCAGCTCCAGAGTCCACCGAGAACTAGGGCG 314
 QY LeuMetAspGlnThrMetTyrGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
 Db 315 CTGATGACGACGACCATGAAAGAGTGAAGCCTACAAATCGAAGAACTGGAGAACAACTG 374
 QY 101 ThrProValAlaGlnGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
 Db 375 ACCCGGGTGGGAGAGAGACGGGGCAAGCTGTCCAAAGAGCTGCAGGGCGCCAGAGCC 434
 QY ArgLeuGlnAlaAspMetGlnAspValCysGlnArgLeuValGlnTyrArgGlnVal 140
 Db 435 CGGCTGGGCGCGGACATGAGAGAGAGTGTGGCGCCCTGTGTGCACTACCGGGGAGAGTG 494
 QY 141 GlnAlaMetLeuAspGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 495 CAGGCGCATGCTCGCGCAGAGACCCAGAGAGCTGGGGTGGCGCTCCCTCCACCTGCGC 554
 QY 161 LysLeuArgLysArgLeuLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 Db 555 AAGCTGGGTAAAGCGGCTCTCCGCGATGCCATGCAAGAAAGCCCTGGCAGTGTAC 614
 QY 181 GlnAlaGlnAlaArgGlnGlnGlnAlaGlnArgGlnLeuSerAlaLeuArgGlnVal 200
 Db 615 CAGGCGCGGGGCGCGGAGAGGGCGCGGAGCGGCTCAGCGCCATCCGAGAGCGGCTGGGG 674
 QY 201 ProLeuValGlnGlnGlnGlnArgValArgAlaAlaThrValGlnSerLeuAlaGlnPro 220
 Db 675 CCCCTGTGTGGAACGGGCGCGGCTGGCGGCCCACTGTGGGTCTCCCTGGCGCGGACGCG 734
 QY 221 LeuGlnGlnLysArgAlaGlnAlaThrGlnGlnArgLeuArgAlaArg-MetGlu--GluMe 239
 Db 735 CTACAGAGAGGGGCGCCAGGCTGTGGGAGAGGGCTGGCGCGGATGGAGGCACATTG 794
 QY 239 TgLYSerArgThrArgAspArgLeuAsp-GluValLysGlnGlnAlaGlnValArgAla 259
 Db 795 GGGCAGCGCGGACCGCGGAGCGCTGTGAACGAGTGAAGAGAGAGTGGAGAAATGGTGGCC 854
 QY 259 LalaLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 279
 Db 855 GCCAAGGCTGGGGGAGACAGGCGCCGCAATTCGCTTGACAGGCCAA-GCCCTTCAGGCGCGCT 913
 QY 279 eulYSerTrpPheGlu 284
 Db 914 CCAAGAGCTGTCTCAA 930
 RESULT 4
 Bg761746 938 bp mRNA linear EST 15-MAY-2001
 LOCUS 602717942F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:481411 5'
 DEFINITION mRNA sequence.
 ACCESSION Bg761746
 VERSION Bg761746.1 GI:14072399
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (Bases 1 to 938)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC/DCTD/DRP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LLCM674 row: c column: 12
 High quality sequence stop: 767.
 Location/Qualifiers
 FEATURES
 source
 1. 938
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="IMAGE:481411"
 /clone_lib="NIH_MGC_49"
 /tissue_type="melanotic melanoma, high MDR (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using Zap-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."

BASE COUNT 171 a 272 c 374 g 121 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.8e-117 Length: 938
 Score: 1205.00 Matches: 270
 Percent Similarity: 91.95% Conservative: 4
 Best Local Similarity: 90.60% Mismatches: 13
 Query Match: 75.83% Indels: 11
 DB: 12 Gaps: 2

US-09-827-854-17 (1-317) x Bg761746 (1-938)
 QY 1 MetLysValLeuThrPpaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 57 ATGAAGGTTCTGTGGGCTGT 116
 QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnTyrGlnSer 40
 Db 117 GAGCAAGCGGTGGAGAGACAGAGCGGAGCCGAGCTGGCCAGCAGACCGAGTGGCAGAGC 176
 QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 60
 Db 177 GGCAGGCGTGGGAAACGCGACCTGGCTCTTTGGATTACCTGGCTGGGTGGCAGACA 236
 QY 61 LeuSerGlnGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
 Db 237 CTGTCTGACGAGGTGCAGAGAGAGCTGTCTCAGCTCCAGGTCAACCGAACTGAGGGCG 296
 QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnGlnLeu 100
 Db 297 CTGATGACGACGACCATGAGAGAGTGAAGGCTTACAAATCGGAAGTGGAGAACTG 356
 QY 101 ThrProValAlaGlnGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
 Db 357 ACCCGGTGGCGGAGAGAGAGCGGGCAGCGCTGTCCAAAGAGCTGACAGGGCGGAGGCC 416
 QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlnArgLeuValGlnTyrArgGlnVal 140
 Db 417 CGGCTGGGCGCGGACATGAGAGAGCTGTGGCGCCCTGTGTGCACTACCGGGGAGAGTG 476
 QY 141 GlnAlaMetLeuAspGlnSerThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
 Db 477 CAGGCGCATGCTGGCCAGAGACCGAGAGAGTGTGGGTGGCTCCCTCCACCTGCGCG 536
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 Db 537 AAGCTGGTAAAGCGGCTCTCCGCGATGCCGATGACCTGACAGAAAGCGCTGGCAGGTAC 596
 QY 181 GlnAlaGlnAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
 Db 597 CAGGCGGGGCGCGCAGAGGCGCGGAGCGGTCTCAGGCGCCATCCGCGAGCGCTGGGG 656

QY 201 ProLeuValGluGlnGlyArgValAlaAlaThrValGlySer-LeuAlaGlyGlnPr 220
|||||
Db 657 CCCCTGTGGAACAGAGCGCGCTGCGGGCCGCCACTGTGGCTCCCTTGCGCGGACGCC 716
QY 220 OleuGlnGluArgAlaGlnAlaTrp-GlyGluArg-LeuArgAlaArgMetGlu---Glu 238
|||||
Db 717 GCTACAGGAGCGGCGCCAGCGCTGGGGGAGCGGCTTGGCGCGGATGAGAGGACAT 776
QY 239 MetGlySerArgThr-Arg-AspArgLeuAspGluValLys-GluGlnValAla---Glu 256
|||||
Db 777 TGGGCAAGCCCGAGCCCGCGACCGCTTGGACGAGTGAGAGGAGCAGGTGGCGGAGC 836
QY 257 ValArgAlaLysLeuGlnGlu-GlnAlaGlnGlnLeuArgLeuGlnAla--GluAlaPhe 275
|||||
Db 837 TTGGGGGCCAAGCTGAGAGACAGGCGCCAGAGATACGCCCTGCAGCGCGGAGAGGCTTC 896
QY 276 GlnAlaArgLeuLysSerTrpPheGluProLeuVal 287
|||||
Db 897 CCAGCGCGGCTCAAAAGCTGTTCGAAAGCCCTGT 932

RESULT 5
B1600906 942 bp mRNA linear EST 07-SEP-2001
LOCUS 603249241P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5301010 5',
DEFINITION mRNA sequence.
ACCESSION B1600906
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 942)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rrmail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshitsugu and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Place: LHAM11762 row: a column: 11
High quality sequence stop: 762.
Location/Qualifiers
1..942
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5301010"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to ROP 5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 177 a 284 c 361 g 120 t
ORIGIN

Alignment Scores:
Pred. No.: 5,18e-117 Length: 942
Score: 1202.50 Matches: 264
Percent Similarity: 92.78% Conservative: 6

Best Local Similarity: 90.72% Mismatches: 15
Query Match: 75.68% Inlets: 8
DB: 13 Gaps: 1
US-09-827-854-17 (1-317) x B1600906 (1-942)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 73 ATGAAGGTTCTGTGGGCTCGTTGCTGTCATATTCCTGCGAGATGCCAGGCCAAGGTG 132
QY 21 GluGlnAlaValGluThrLupProGluProGluLeuArgGlnGlnThrLupTrpGlnSer 40
|||||
Db 133 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCGCACAGACAGCAGTGGCAGAGC 192
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLysArgTrpValGlnThr 60
|||||
Db 193 GGCACGCTGGGAACTGGCTGCTGCTGCTTTTGGATTTACTGCTGCTGGGTGCAGACA 252
QY 61 LeuSerGluGlnValGlnGluLeuLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
Db 253 CTGTCTGACACAGTGCAGAGAGAGCTGCTACACTCCAGTCCACAGAGAACTGAGGCGC 312
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGluGlnLeu 100
|||||
Db 313 CTGATGGAGACAGACCATGAGAGTGAAGCGCTTCAAAATCGAACTGGAGAGAACACTG 372
QY 101 ThrProValAlaGluGluThrArgAlaArgLysSerLysGluGlnAlaGlnAla 120
|||||
Db 373 ACCCGGTGGCGGAGAGAGCGCGGCGCTGTCCAGAGACTGAGGCGGCGAGGCGC 432
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnThrArgGlyGluVal 140
|||||
Db 433 CGGCTGGGGCGGAGACATGAGAGACGCTGCGCGCGCTGTGTCAGTACCGCGGCGAGTGC 492
QY 141 GlnAlaMetLeuAspGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 493 CAGGCCATGCTCGCGGACAGACCGGAGCTGCGGCTGCGCTGCCCTCCACCTCGCGC 552
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
|||||
Db 553 AAGCTGCGTACCGGCTCTCCGCGATGCGATGACCTGCAAGAGCGCTTGGCAGTGTAC 612
QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaTrpArgGluArgLeuGly 200
|||||
Db 613 CAGGCGGGGCGCGGAGGCGCGGAGCGCGGCGCTCAGCGCATCCGAGGCGCTCGGGG 672
QY 201 ProLeuValGluGlnGlyArgValAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
|||||
Db 673 CCCCTGTGGAACAGAGCGCGCTGCGGCGC-GCCACTGTGGCTCCCTGCGCGGAGCGC 731
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAla-ArgMetGlu-GluMetG 240
|||||
Db 732 CTACAGAGAGCGGCGCGAGGCTGTGGGCGAGCGGAGGCGCGGAGTGGAGGAGATTG 791
QY 240 LysSerArgThrArgAspArgLeuAsp-GluValLys-GluGlnValAlaGluVal-Arg 259
|||||
Db 792 GAGGCGGAGCGCGGAGCGCTGAGACGAGGAGACGAGGTTGCGGAATGTCGCGG 851
QY 255 LysLeuGluGluGlnAla---GlnGlnTrpArgLeuGlnAlaGluAlaPheGlnAla 278
|||||
Db 852 CCACCTTGGAGGAGACAGCGCCGAGCATATACGCTTGGAGGCGGAGGCTTCAGAGGCC 911
QY 278 rGleuLysSerTrpPheGluPro 285
|||||
Db 912 GCTCTAAAC-TGTTGACCCCT 933

RESULT 6
BM042094 800 bp mRNA linear EST 07-NOV-2001
LOCUS BM042094
DEFINITION 603615713P1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5420618 5',
ACCESSION mRNA sequence.
BM042094
VERSION BM042094.1 GI:16771361

Alignment Scores:

Pred. No.: 1,97e-115 Length: 927
 Score: 1187.50 Matches: 273
 Percent Similarity: 91.67% Conservative: 2
 Best Local Similarity: 91.00% Mismatches: 15
 Query Match: 74.73% Indels: 12
 DB: 12 Gaps: 1

US-09-827-854-17 (1-317) x BG472299 (1-927)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB ATGAAGGTTCTGTGGGCTCGTTGCTGTGCATCTCTGCGAGATGCCAGGCCAAGGTG 106
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB GAGCAAGCGGTGGAGACAGACGCGGAGCCGAGCTGCGCACAGACCGAGTGGCGAGAC 166
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB GGCCACGCTGGGAACCTGGCAGCTGCTTTGGATTACCTGCGGTGGGTGGAGACA 226
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB CTGTCTGACAGGTGACAGAGGAGGTGCTCAGCTCCAGCTCACCGAGAACTGAGAGGCG 286
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100
 DB CTGATGACAGACGACCATGTAAGGAGTTGAAGGCTTCAAAATCGAACTGAGAGAACTAG 346
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB ACCCGGTGGCGGAGAGACGGGGCAGCGGCTGTCAGAGACTGCGAGCGCGCGAGGC 406
 QY 121 ArgLeuGlyAlaAspMetLysAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 DB CGGCTGGGCGCGACATGAGAGACGTGTGCGGCGCTGTGTCAGTACCGCGGCGAGGTG 466
 QY 141 GlnAlaMetLeuAspGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB CAGGCGATCTCGCGCCAGAGCACCGAGGAGCTGGGGTGGCTGCCCTCCACCTGCGCG 526
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 DB AAGCTCGTAAGCGGCTCTCCGCGATGCCGATGACTCCAGAAAGCGCTTGGCAGGTAC 586
 QY 181 Gln-AlaGlyAlaArgGluGlyAlaGluArgGly--LeuSerAlaIleArgGluArgLeu 199
 DB CAGGCGCGGGGCGCGGAGGCGCGCGCCCTCAAGGCCCATCCGCGCGAGCCCTT 646
 QY 200 --GlyProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyG 219
 DB GGGGCGGCTGGGTGGAAACAGGCGCGGTGCGGCGCCGCTGTGGCTCTCCCTGGCGGCC 706
 QY 219 LnpProLeuGlnGluArgAlaGlnAlaTrpGlyLysArgLeuArgAlaArgMetGluGln 239
 DB AAGC-CTACAGAGCGCGCGCCAGGCTGGGCGGAACGCTG-CGCGCGCGAGTGGAGAGA 764
 QY 239 eegLysSerArg-ThrArgAspArgLeuAspGluValLysGluGlnValAlaGlu-ValAr 258
 DB TGGGGGAGGGGAGCCGCGAGCTGTGACCGAGGTGAAGACAGAGTGGCGAGGGTGGG 824
 QY 258 gAlaLysLeuGluGlnGluAlaGlnGlnIleArgLeuGlnAlaGlnAla---PheGlnAl 277
 DB CGCAAGCTGGAGGAAGCAAGCGCAGATACGAGCTGCGAGCGCGAGGGGCTTCCAGGGC 884
 QY 277 aArg-LeuLysSerTrpPhe-GluProLeuValGluLysMet 290
 DB CCGGCTCAAGAGCTGTTCGAGAGCCCTGGGTGAAGACATG 926
 RESULT 8
 B1597743 LOCUS B1597743 922 bp mRNA linear EST 07-SEP-2001

DEFINITION 603248609f1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5300259 5',
 mRNA sequence.
 ACCESSION B1597743
 VERSION B1597743.1 GI:15490682
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 922)
 AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLM11760 row: b column: 04
 High quality sequence stop: 782.
 Location/Qualifiers
 1..922
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5300259"
 /clone_1ib="NIH_MGC_96"
 /tissue.type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: Brain; Vector: Bluescript (modified
 Bluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3',
 size-selected for average insert size 2.3 kb and
 normalized to ROP 5. This is a primary library enriched
 for full-length clones and constructed using the
 cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
 BASE COUNT 176 a 273 c 359 g 114 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,05e-115 Length: 922
 Score: 1184.50 Matches: 255
 Percent Similarity: 92.83% Conservative: 4
 Best Local Similarity: 91.40% Mismatches: 16
 Query Match: 74.54% Indels: 5
 DB: 13 Gaps: 1
 US-09-827-854-17 (1-317) x B1597743 (1-922)
 QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB ATGAAGGTTCTGTGGGCTCGTTGCTGTGCATCTCTGCGAGATGCCAGGCCAAGGTG 132
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB GAGCAAGCGGTGGAGACAGACGCGGAGCCGAGCTGCGCACAGACCGAGTGGCGAGAC 192
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB GGCCACGCTGGGAACCTGGCAGCTGCTTTGGATTACCTGCGGTGGGTGGAGACA 252
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB CTGTCTGACAGGTGACAGAGAGCTGCTCAGCTCCAGAGTACCCAGAACTGAGAGGCG 312
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100


```
Db 313 CTGATGACGACACCATGAAGAGTTGAGGCTTACAAATCGCACTGGAGGACAACTG 372
Qy 101 ThrProValAlaIaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaGlnAla 120
Db 373 ACCCCGGTGGCGAGGAGACGGGGGACGGCTGTCCAGAGAGCTGCAGGGGGGCGACAGCC 432
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 433 CGGCTGGCGCCGACATGAGGAGCGTGTGGCGCGCTGTGACGTACCCCGCGGAGGTG 492
Qy 141 GlnAlaMetLeuAspGlnSerThrGluGluLeuValArgLeuValAspSerHisLeuArg 160
Db 493 CAGGCCATGCTCGGCCACGAGACCGAGAGAGCTGGGGTGGCTCGCTCCGCTCCACCTGGCG 552
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 553 AACCTGGGTAAAGGGGCTCTCCGCGATGCCGATGACCTGCAGAAAGCGCCCTGGCAGTGTAC 612
Qy 181 GlnAlaGlyAlaArgGluGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 613 CAGGCGCGGCGCCGCGAGGGCGCGGAGCGCGGCTC-AGCCCATCCGCGAGCGCCCTGGGG 671
Qy 201 ProLeuVal-GluGlnGlyArgValArgAlaIleThrValGlySerLeuAlaGlyGln-P 220
Db 672 CCCCTGGTGGGAACAGGGGCGGGTCCGCGCCACTGTGGCTCCCTGGCGGCGACGCC 731
Qy 220 roLeuGlnGluArg-AlaGlnAlaTyrPglYgluArgLeuArgAlaArgMetGluGluMet 239
Db 732 CGGTACAGAGAGGGGGGCGCCAGGCTGGGGGAGCGGATACGGCGCGATGGAGAGAGAG 791
Qy 240 GlySerArgThrArgAspArgLeuAspGlyValLysGluGlnValAlaGluValArgAla 259
Db 792 GGCAGCGCGGACCCGAGACCCCTGCGACGAGGTAAAGAGCAGGTGGCGGGGCGGCGCA 851
Qy 260 LysLeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGln 276
Db 852 CAAGCTGAAGAAACAGCCCAAAATACGCTTA---GCCGAGAGCTTCAGG 899

RESULT 9
Bg774871 817 bp mRNA linear EST 15-MAY-2001
LOCUS Bg774871
DEFINITION 602649975F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4760827 5',
mRNA sequence.
ACCESSION Bg774871
VERSION Bg774871.1 GI:14045188
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 817)
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: DCTD/DP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LICM1612 row: e column: 20
High quality sequence stop: 813.
Location/Qualifiers
1..817
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4760827"
/clone_lib="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
```

```
/note="Organ: prostate; Vector: pOTB7; Site_1: xhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
```

```
BASE COUNT 148 a 242 c 324 g 103 t
ORIGIN
Alignment Scores:
Pred. No.: 3,06e-114 Length: 817
Score: 1175.50 Matches: 250
Percent Similarity: 96.55% Conservative: 2
Best Local Similarity: 95.79% Mismatches: 7
Query Match: 73.98% Indels: 4
DB: 12 Gaps: 0
US-09-827-854-17 (1-317) x Bg774871 (1-817)
Qy 1 MetLysValLeuThrPalaIaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 40 ATGAGAGTTCTGTGGGCTGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 99
Qy 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluThrGlnSer 40
Db 100 GAGCAAGCGGTGGAGACAGAGCCGAGAGCCGAGCTCCGCGACAGACAGACAGAGAGAGAG 159
Qy 41 GlyGlnArgTyrPglLeuLeuAlaLeuGlyArgPheTyrAspTyrLeuArgTyrPalaGlnThr 60
Db 160 GGCACAGCGCTGGGAACGTGGACACTGGGCTCTTTGGATTACCTGGCGGGGTGGAGACA 219
Qy 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 220 CTGTCTGAGCAGGTGCGAGAGAGAGAGTGTCTGAGTCTCCAGGTACCCAGAGAGAGAGGCG 279
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGluLeu 100
Db 280 CTGATGACGAGACCATGAAGAGGTGAAGGCTTACAAATCGAATCGAGAGCAACACTG 339
Qy 101 ThrProValAlaIaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaGlnAla 120
Db 340 ACCCCGGTGGCGAGGAGACGGGGGCGCGCTGTCCAGAGAGCTGAGCGGGCGCGAGGCC 399
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 400 CGGCTGGCGCGGACATGAGAGAGCTGTGCGCGCGCTGTGTCAGTACCGCGCGAGGTG 459
Qy 141 GlnAlaMetLeuAspGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 460 CAGGCCATGCTCGGCCAGAGACCCGAGAGCTGCGGCTCGCTCCCTCCACACCTGGCC 519
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 520 AAGCTTCGTAAAGGGGCTCTCCGCGATGCCGATCTCCAGAAAGCGCTGGGAGAGTAC 579
Qy 181 GlnAlaGlyAlaArgGluGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 580 CAGGCCGGGGGCGCGAGGAGCGCGAGCGCGCTAGAGCCATCCGCGAGGCGCTGGGG 639
Qy 201 ProLeuValGluGlnGlyArgValArgAlaIleThrValGlySerLeuAlaGlyGlnPro 220
Db 640 CCCCTGGTGAAGAGGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 698
Qy 221 LeuGlnGluArgAlaGlnAla--TyrPglYgluArgLeuArgAlaArgMetGluGluMetG 240
Db 699 CTACAGAGAGCGGGCCGAGGCTTTGGGGAGCGAGTGGCGCGCGCGAGTGGAGAGATGG 758
Qy 240 LysSerArgThrArgAspArgLeuAspGlyValLysGluGlnValAlaGluValArgAla 259
Db 759 GCCACCGGAGCG--GAGCGCTGGAGAGGTGAAGAGAGAGGTGGCGGAGGTGGCGGCC 815
```


RESULT 10
 BI600563
 LOCUS 811 bp mRNA linear EST 07-SEP-2001
 DEFINITION 603244936p1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:5287329 5',
 mRNA sequence.
 ACCESSION BI600563
 VERSION BI600563.1 GI:15493502
 KEYWORDS EST.
 SOURCE human.
 ORGANISM human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 811)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ULNL at:
 http://image.llnl.gov
 Plate: LHAM11726 row: 9 column: 10
 High quality sequence stop: 783.
 Location/Qualifiers
 1..811
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5287329"
 /clone_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescript (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgaag
); Oligo-dT primed using primer 5'-TTTGTGTGTGTGTGT-3',
 size selected for average insert size 2.3 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the
 cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
 BASE COUNT 148 a 245 c 316 g 102 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4.37e-114 Length: 811
 Score: 1174.00 Matches: 239
 Percent Similarity: 97.56% Conservative: 1
 Best Local Similarity: 97.15% Mismatches: 5
 Query Match: 73.88% Indels: 1
 DB: 13 Gaps: 0
 US-09-827-854-17 (1-317) x BI600563 (1-811)
 QY 1 MettysValleuTPAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||||||
 DB 74 ATGAAGGTTCTGTGGCTCTGCTTGGCTGCACATTCCTGGCAGATGCCAGGCCAAGGTG 133
 |||||||
 QY 21 GtuglnAlaValGluThrcGluProGluProGluLeuArgGlnGlnThrcGluTrpGlnSer 40
 |||||||
 DB 134 GAGCAAGCGGTGGAGACACAGCCGAGCTGGCCAGCAGACGACGAGGCGAGAC 193
 |||||||
 QY 41 GtlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 |||||||
 DB 194 GGCACAGCGCTGGGAACCTGCACTGGTGGCTTTGGGATTACCTGGCGGTGGCAGAGA 253
 |||||||
 QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 |||||||
 DB 254 CMTGTGACAGGTGACGAGAGAGCTGCTCACCTCCAGCTCACCAGCACTGAGGGCG 313
 |||||||

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 |||||||
 DB 314 CMTGTGACAGGTGACGAGAGAGCTGGAAGGCTCAAAATCGGAAGTGGAGAGACACAG 373
 |||||||
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
 |||||||
 DB 374 ACCCGGTGGCGAGAGAGCGCGGCTGTCCAGAGACTCCAGCGCGCGCAGGCC 433
 |||||||
 QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnThrArgGlyGlnVal 140
 |||||||
 DB 434 CGGCTGGGGCGGACATGAGAGACGTGTGGCCGCTGTGTACAGTACCGCGCGAGGTG 493
 |||||||
 QY 141 GlnAlaMetLeuAspGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||||
 DB 494 CAGCGCATCTCTGGCCAGACGACACCGAGAGCTGGGGTGGCTGGCTCCACCTCCGCC 553
 |||||||
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 |||||||
 DB 554 AAGCTGCGAAGCGGCTCTCCGATGCGCATGACCTCAGAGAGCGCTCGACAGTAC 613
 |||||||
 QY 181 GlnAlaGlyAlaArgGlnGluGluArgGlyLysSerAlaIleArgGluArgLeuGln 200
 |||||||
 DB 614 CAGGCGGGGGCCGCGAGGGCGCCGAGCGGCTTCAGGCGCATCCGAGCGCGCTGGG 673
 |||||||
 QY 200 ProLeuValGluGlnGlyArgValArgAlaIleThrValGlySerLeuAlaGlyGlnPr 220
 |||||||
 DB 674 CCCCTTGGTGAACAGAGCGCGGCTGGGGCCGCCACTGTGGCTCCCTGCCAGCAGCGC 733
 |||||||
 QY 220 OLeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGlnMetG1 240
 |||||||
 DB 734 GCTACAGAGAGCGGGCCAGAGCTGGGGCGAGCGCTGCGCGCGGATGAGAGCATGGG 793
 |||||||
 QY 240 ySerArgThrArgAsp 245
 |||||||
 DB 794 CAGCGGAGACCGCGAA 809
 |||||||
 RESULT 11
 BG829472
 LOCUS 845 bp mRNA linear EST 22-MAY-2001
 DEFINITION 602763768p1 NIH_MGC_42 Homo sapiens CDNA clone IMAGE:4899112 5',
 mRNA sequence.
 ACCESSION BG829472
 VERSION BG829472.1 GI:14177059
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 845)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ULNL at:
 http://image.llnl.gov
 Plate: LCM1790 row: 9 column: 17
 High quality sequence stop: 829.
 Location/Qualifiers
 1..845
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="4899112"
 /clone_lib="NIH_MGC_42"
 /tissue_type="epitheloid carcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: pancreas; Vector: pOTB7, Site_1: XhoI;

QY	61	1eserGluGlnValGlnGluGlnLeuLeuSerSerGlnValThrGlnGlnLeuLeuArg1a	80
Db	253	CTGTTGACACAGTCCACAGAGAGCTGCTCAGCTCCACAGTCCACAGAGAACTGAGGGCG	312
QY	81	LeuMetAspGluThrMetIysGluLeuLeuLysAlaTyrIysSerGlnLeuGlnGlnLeu	100
Db	313	CTGATGCACGAGACCATGAGAGAGTTGAAGGCTTCACAAATGCGAATCGAGAGCAACTG	372
QY	101	ThrProValAlaGluGluThrArgAlaArgLeuSerIysGluLeuGlnAlaAlaGlnAla	120
Db	373	ACCCGGCTGGGGGAGACACGCGGCGACGCGTGTCCAAAGAGAGCTTCAGCGGCGCAGGCC	432
QY	121	ArgLeuGluValAspMetGluAspPheLysGlyArgLeuValGlnTyrArgIysGluVal	140
Db	433	CGGCTGGGCGCGGACATGAGAGACGTGTGGCGGCCCTGTGTGACGTACCGCGGAGGTG	492
QY	141	GlnAlaMetLeuAspGlnSerThrGluGlnLeuLeuArgValAlaGlnLeuAlaSerHisLeuArg	160
Db	493	CAGGCAATGCTCGGCGCAGAGCACCGAGGAGGTGGGGTGGCGCTCGCTCCACCTGGGCG	552
QY	161	LysLeuArgIysArgLeuLeuArgAspAlaAspAspLeuGlnIlyCysLeuAlaValIlyr	180
Db	553	AAGCGCTTAAGCGGCTCTCCGCGCATCCGATCCGACGACGAGAACGCGCTGCGCATGTAC	612
QY	181	GlnAlaGluValArgGluGlnGluValaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200
Db	613	CAGCGCGGGGCGCGCGAGAGCGCCCGACGCGCGGTCTCACCCCATCCGCGAGCGCTGGGG	672
QY	201	ProLeuValGluGlnGlyArgValAlaArgAlaAlaThrValGlySerLeuAlaGlyGlnPro	220
Db	673	CCCCGTGTGAACAAGGCGCGGTGGCGGCGCGCACTGTGGGCTCCCTGGCCAGCGACGCG	732
QY	221	LeuGlnGluArgAlaGlnAlaThrPheGlyLeuArgLeuArgAlaArgMetGluGlnMet	239
Db	733	CTACAGAGGCGGCGCCAGGCTGGGGCGCAGCGGTG-CGCGCGCGGATGAGAGATG	788
RESULT	13		
LOCUS	B1551475	919 bp	mRNA linear EST 05-SEP-2001
DEFINITION	60319431F1 NIH_MGC_95	Homo sapiens	CDNA clone IMAGE:5274003 5',
ACCESSION	B1551475		
VERSION	B1551475.1	GI:15438787	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Exxarvota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 919)		
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapds-remail.nih.gov		
	Tissue Procurement: Miklos Palkovits, M.D., Ph.D.		
	CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shizaki		
	Toshiyuki and Piero Carninci (RIKEN)		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC Clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNLN at:		
	http://image.llnl.gov		
	Plate: L1AM11691 row: 1 column: 04		
	High quality sequence stop: 812.		
FEATURES	Location/Qualifiers		
SOURCE	1..919		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5274003"		
	/clone_lib="NIH_MGC_95"		
	/tissue_type="hippocampus"		
	/lab_host="DH10B"		

BASE COUNT	172 a	270 c	363 g	114 t
ORIGIN	/note="Organ: brain; Vector: pBluescriptPR (modified pBluescript KS+). Site_1: BamH; Site_2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-ttttttttttttt-3', size-selected for average insert size 2.5 kb and normalized to ROP 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH-MGC Library."			
Alignment Scores:	9.01e-110	Length:	919	
Pred. No.:	1134.00	Matches:	245	
Score:	92.88%	Conservative:	3	
Percent Similarity:	91.76%	Mismatches:	18	
Best Local Similarity:	71.37%	Indels:	4	
Query Match:	13	Gaps:	0	
DB:				
US-09-827-854-17 (1-317) x BI551475 (1-919)				
QY	1	McttysValleuTTPAlaAlaIleuLeuValThrPheLeuAlaGlyCysGlnAlaIaLysVal	20	
Db	73	ATGAAGATTCTTGGGGCTGGCTGGTGGTGCATTTCTGGCAGGATGCCAGGCCAGGTG	132	
QY	21	GtuglnAlaIaValGluThrGluProGluIleuLeuAArgGlnGlnThrGluTTPInser	40	
Db	133	GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGGCCAGACAGACCGAGTGGCAGAGC	192	
QY	41	GlyGlnAArgTTPGluLeuAlaLeuGlnIaArgPheTrpAspTyrLeuAArgTTPValGlnThr	60	
Db	193	GGCCAGCGCTGGGAAGTGGCACTGGCTGGTGGCTTTGGATTACCTGGCTGGGTGCAGACA	252	
QY	61	LeuSerGluGlnValGlnGluIleuLeuSerSerGlnValThrGlnGluLeuArgIa	80	
Db	253	CTGTCTGAGCAGTGCAGAGAGAGCTGCTCACTCCAGGTCACCCAGGAAGCTGAGGGCG	312	
QY	81	LeuMeTAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu	100	
Db	313	CTGATGAGACAGACCATGAGAGATTGAAGGCTTCAATTCGGAATCGAGAACAACTG	372	
QY	101	ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaGlnAla	120	
Db	373	ACCCGGTGGCGGAGAGACAGCGGCGACGGCTGTCCAAGAGATCCAGCGGCGAGGCC	432	
QY	121	ArgLeuGlnIaIaAspMetGluAspAlaLysGlyAArgLeuValGlnTyrArgGlyVal	140	
Db	433	CGGCTGGCGCGCATGTAGAGACTGTGGCGCCCTGGTGCATTCACGGCGCAGGTG	492	
QY	141	GlnAlaMetLeuAspGlnSerThrGluGluLeuAArgValAArgLeuAlaSerHisLeuArg	160	
Db	493	CAGGCCATGTCTGGCCAGACACCGAGAGAGCTGGCGGTGCCCTGCCCTCCACCTGGCC	552	
QY	161	LysLeuAArgLysAArgLeuLeuAArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr	180	
Db	553	AAGTCGCGTAAGCGGCTCTCCGCGATGCCGATGCCATCGCAAGAGCGCTGCAGGTAC	612	
QY	181	GlnAlaGlyAlaIaArgGlnGlyAlaGluArgGlyLeuSerAlaIleAArgGlnAArgLeuGly	200	
Db	613	CAGCGGGGGGGCGGAGTGGCGCCAGGCGGTGCTCACAGGCATCCGGAGACGCTGGGG	672	
QY	201	ProLeuValGlnGlnGlyArgValAArgAlaIaThrValGlySerLeuAlaGlyGlnPro	220	
Db	673	CCCTGTGTGGAAAGAGGGCGCGTGGGGCCGCCACTGTGTGGGTCCCTGGGGCGCAGCCCT	732	
QY	221	LeuGlnGlnAArgAlaGlnAlaTTPGlyGlnAArgLeuAArgAlaArgMetGluGlu-MetGln	240	
Db	733	A--CAGGAGCGGGCGCAGGCTGTGGGGCGAGGGGTCCGCGCGGATGAGAGAAATGGG	790	
QY	240	ySerAArgThrAArgAspAArgLeuAspGluValLysGlnValAlaGluValAArgAlaIaIy	260	
Db	791	CAGCGGAGCCGGGGACG--CTGGAGGAATGAAAGACAGCTGGGGGGGAGG--TGGCCCA	847	

OY 260 sleuglgluglalnagl 266
|||||:::||||
Db 848 GCTGAGAGACGACGACAG 866

RESULT 14
BM042228
LOCUS 757 bp mRNA linear EST 07-NOV-2001
DEFINITION 60361616F1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5557004 5',
mRNA sequence.
ACCESSION BM042228
VERSION BM042228.1 GI:16771495
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 757)
AUTHORS NIH-MGC http://mgi.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@rs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2021 row: k column: 21
High quality sequence stop: 757.

FEATURES
source

1. 757
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5557004"
/clone_lib="NIH_MGC_112"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOT7; Site: 1: XhoI; Site: 2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."

BASE COUNT 134 a 229 c 256 g 98 t
ORIGIN

Alignment Scores:
Pred. No.: 7.83e-109 Length: 757
Score: 1124.00 Matches: 233
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 3
Query Match: 70.74% Indels: 2
Gaps: 0

US-09-827-854-17 (1-317) x BM042228 (1-757)

OY 1 MetlYValleuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
Db 50 ATGAAGATTCTGTGGCTGCTGTCATTCCTGCGCAGATGCCAGGCAAGTG 109
|||||
OY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluThrGlnSer 40
|||||
Db 110 GACCAAGCGGTGAGACAGCGCGAGCCGACCTCGCAGCAGACCGAGTGGCAGAGC 169
|||||
OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
Db 170 GCGCAGCGCTGGAACTGCACCTGGCTGTTTGGGATTACTGCGCTGGGTGCAGACA 229
|||||

OY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
Db 230 CTGCTTACGACAGGTGCGAGAGAGCTGCTCAGCTCCAGCTCACCAGAACTGAGGAGCTG 289
|||||
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
|||||
Db 290 CTGATGACAGACACCATGAAAGAGTTGAAAGCCCTACAAATCGAATCGAGGAGAACACTG 349
|||||
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
|||||
Db 350 ACCCGGTGCGAGAGAGAGCGGGCAGCGCTGTCCAGAGAGCTGACAGCGCGCGAGCGCC 409
|||||
OY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
|||||
Db 410 CGGCTGGCGCGGACATGAGGAGCGTGTGGCGCGCTGTGTCAGTACCCCGCGAGGTG 469
|||||
OY 141 GlnAlaMetLeuAspGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 470 GAGGCAATGCTCGCCACAGACACCGAGACTCGGGTGGCTCCGCTCCACCTGCGC 529
|||||
OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
|||||
Db 530 AAGCTGCTAAGCGGCTCCCTCCGATGCCGATGACCTGCAGAAAGCGCTGCGCAGTGTAC 589
|||||
OY 181 GlnAlaGlyAlaArgGluGluGlyAlaGluArgGlyLeuSerAlaThrArgGluArgLeuGly 200
|||||
Db 590 CAGCGCGGCG 649
|||||
OY 201 ProLeuValGluGlnGlyArgVal-ArgAlaAlaThrValGlySerLeuAlaGlyGlnP 220
|||||
Db 650 CCCTGTGTGAAACAGGCG 709
|||||

RESULT 15
Bg706129 907 bp mRNA linear EST 07-MAY-2001
LOCUS 60266903F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4792030 5',
mRNA sequence.
DEFINITION Bg706129 GI:13981169

ACCESSION Bg706129
VERSION Bg706129.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 907)
AUTHORS NIH-MGC http://mgi.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@rs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
Toshinuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM10669 row: 1 column: 23
High quality sequence stop: 832.

FEATURES
source

1. 907
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4792030"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptP (modified)

pbuescript K5+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROR 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carnucci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 171 a 260 c 358 g 117 t 1 others
ORIGIN

Alignment Scores:

Pred. No.:	1.03e-107	Length:	907
Score:	1114.50	Matches:	242
Percent Similarity:	93.85%	Conservative:	2
Best Local Similarity:	93.08%	Mismatches:	12
Query Match:	70.14%	Indels:	6
DB:	12	Gaps:	2

US-09-827-854-17 (1-317) x BG706129 (1-907)

```
OY      1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db      |||||||
OY      75 ATGAAGTTCTGTGGGCTGCTGTTGCTGTCACATTCTGGCAGATGCCAGGCCAAGGTG 134
Db      |||||||
OY      21 GlnGlnAlaValAluThrGlnProGlnProGlnLeuArgGlnGlnThrGlnTrpGlnSer 40
Db      |||||||
OY      135 GAGCAAGCGGTGAGACAGACGCCGAGCCGAGCTGCCAGCAGACCGAGTGGCAGAC 194
Db      |||||||
OY      41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
Db      |||||||
OY      195 GGCAGCGCTGGGAAGTGGCTGCTTTGGGATTACCTGGCTGGGTGGCAGACA 254
Db      |||||||
OY      61 LeuSerGlnGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db      |||||||
OY      255 CTGTGTGACAGGTGCAGAGAGAGCTGCTCAGCTCCAGTCACCCAGGAAGTGAAGGGCG 314
Db      |||||||
OY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu 100
Db      |||||||
OY      315 CTGATGAGACGACCATGATGAGAGTTGAAGGCTTCAATTCGAACTGGAGGACACTG 374
Db      |||||||
OY      101 ThrProValAlaGlnGluThrArgAlaArgLeuSerLysGlu-LeuGlnAlaAlaGlnAl 120
Db      |||||||
OY      375 ACCCGGTGGCGAGAGACGGCGACCGCTGCCAAGAGCTGCAGCGCGGCGAGGC 434
Db      |||||||
OY      120 aArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVa 140
Db      |||||||
OY      435 CCGGCTGGGCGCGACATGAGAGACGTGTGCGCGCTGTGTGACGTACCGCGGCGAGGT 494
Db      |||||||
OY      140 LglnAlaMetLeuAspGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuAr 160
Db      |||||||
OY      495 GAGGCGCATGTCTGGCGCAGACACCGAGAGAGTGGCGGCTGCCTCCACCTGCGG 554
Db      |||||||
OY      160 glySLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTr 180
Db      |||||||
OY      555 CAAGCTGCTAGCGGCTCTCCGCGATGCCATGACCTGCAGAAAGCGCTGCGAGTGA 614
Db      |||||||
OY      180 rGlnAlaGlyAlaArgGlnGlyAlaGlnArgGlyLeuSerAlaIleArgGluArgLeu-G 200
Db      |||||||
OY      615 CCAGGCGGGGCGCGGAGGCGCGGAGCGGCGCTCAGCGGCATCCGCGAGCGCTGGG 674
Db      |||||||
OY      200 LyrProLeuValGlnGlnArgValArgValAlaAlaThrValGlySerLeuAlaGlyLnp 220
Db      |||||||
OY      675 GCCCTGTGTGACAGAGGCGCGTGGCGGCCCACTGTGGCTCCCTGCGNCGGCAGC 734
Db      |||||||
OY      220 roLeuGlnGlnArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnLmetG 240
Db      |||||||
OY      735 CGCTACAGAGCGGCGCAGGCTG---GGCGAGCGGTGCGCGCGGATGAGAGATGG 791
Db      |||||||
OY      240 LysArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArg 258
Db      |||||||
OY      792 GCAG--CGGACCCGG---ACGCTGACGAGCTG--AAGAGCAGTGGGGAAGCTGTGCGC 842
Db      |||||||
```

Search completed: March 14, 2003, 20:14:10
Job time : 1296.69 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:08:57 ; Search time 2307.23 Seconds
(without alignments)
3998.351 Million cell updates/sec

Title: US-09-827-854-18

Perfect score: 1589
Sequence: 1 MKVLMALLVFLAGCQAKV.....VEKYQAVGTSAAPVSDNH 317

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+._p2n.model -DEV=xlp
-O=/cgn2.1/USPTO/US09827854/rumat_11032003_101610_27486/app_query.fasta_1.3576
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdl -LIST=45
-OCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=fto -NORM=ext -HEADSIZE=5001 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854.ecgn_1.1.13008.0.rumat_11032003_101610_27486 -NCPU=6 -ICPU=3
-NO_XLPRX -NO_MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database :

GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vl:*
15: gb.da:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vl:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vrl:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1577	99.2	1110	6 E00359	E00359 cDNA coding
2	1577	99.2	1110	6 E00823	E00823 DNA sequence
3	1577	99.2	1147	6 AX302545	AX302545 Sequence
4	1577	99.2	1156	6 BD004278	BD004278 Apo E hum
5	1577	99.2	1156	9 HDMAPOE3	K00396 Homo sapien
6	1577	99.2	1186	9 BC003557	BC003557 Homo sapi
7	1570	98.8	1110	6 E08423	E08423 DNA coding
8	1569	98.7	1156	6 BD004277	BD004277 Apo E hum
9	1568	98.7	1157	6 AX333278	AX333278 Sequence
10	1568	98.7	1157	6 AX409597	AX409597 Sequence
11	1568	98.7	1157	6 I15975	I15975 Sequence 1
12	1568	98.7	1157	6 HDMAPOE	M12529 Human apoli
13	1565	98.5	1156	6 BD004279	BD004279 Apo E hum
14	1443	90.8	1178	9 MFAPOE	X13887 Monkey mRNA
15	1388.5	87.4	5491	9 AF261279	AF261279 Homo sapi
16	1388.5	87.4	41907	6 AX358722	AX358722 Sequence
17	1388.5	87.4	41907	9 AF050154	AF050154 Homo sapi
18	1388.5	87.4	107565	9 AC011481	AC011481 Homo sapi
19	1371	86.3	5515	9 HUMAPOE4	M10065 Human apoli
20	1342	84.5	5413	9 AF261280	AF261280 Pan trogl
21	1259	79.2	4762	9 BABAPOE	M29322 Baboon apol
22	1258	78.7	208239	2 AC021988	AC021988 Homo sapi
23	1250	78.7	1138	4 AF303830	AF303830 Tupaia gl
24	1162.5	73.2	965	6 AX384545	AX384545 Sequence
25	1162.5	73.2	1108	4 BRAPOLPE	X61171 B. taurus mr
26	1162.5	73.2	5617	6 AX384541	AX384541 Sequence
27	1162.5	73.2	6026	6 AX384539	AX384539 Sequence
28	1162	73.1	1060	4 BABAPOLP	M36603 Rabbit apol
29	1144.5	72.0	1154	4 BRAPOLMR	X64839 B. taurus mr
30	1143	71.9	718	9 AF200499	AF200499 Pan trogl
31	1141	71.8	1045	10 MUSAPOE	M12414 Mouse apoli
32	1141	71.8	1104	10 BC028816	BC028816 Mus muscu
33	1139	71.7	718	9 AF200500	AF200500 Gorilla g
34	1139	71.7	718	9 AF200503	AF200503 Pongo pyg
35	1135	71.4	718	9 AF200506	AF200506 Hylobates
36	1135	71.4	1122	4 SSAPOE	X72835 S. scrofa mr
37	1130	71.1	1126	6 AR164342	AR164342 Sequence
38	1130	71.1	1126	6 AR205885	AR205885 Sequence
39	1086	68.3	951	10 S76779	S76779 rAPOE-apol
40	1064	67.0	959	10 MUSAPOE	M73490 Mus musculu
41	1048.5	66.0	1069	10 RATAPOE	J00705 rat apoli
42	1021	64.3	228698	2 AC127479	AC127479 Mus muscu
43	1021	64.3	237653	2 AC073760	AC073760 Mus muscu
44	995	62.6	4856	10 MUSAPE	D00466 Mus musculu
45	994	62.6	4267	6 AR164387	AR164387 Sequence

RESULT 1

ALIGNMENTS

LOCUS	E00359	1110 bp	RNA	linear	PAT 29-SEP-1997
DEFINITION	cDNA coding human apolipoprotein E3.				
ACCESSION	E00359				
VERSION	E00359.1	GI:2168646			
KEYWORDS	JP 1985118189-A/1.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 1110)				
TITLE	Teranishi, Y., Takamitsu, N., Matsui, Y., Kimura, M. and Ikeda, Y.				
JOURNAL	DNA FRAGMENT.				
COMMENT	Patent: JP 1985118189-A 1 25-JUN-1985; MITSUBISHI CHEM IND LTD OS human PN JP 1985118189-A/1 PD 25-JUN-1985 PF 29-NOV-1983 JP 1983224980 PI TERANISHI YUTAKA, TAKAMATSU NOBUHIKO, MATSUI YASUSHI, PI KIMURA MASAKO, PI IKEDA YASUOKO PC C12N15/00, C07H21/04//C12P21/00; CC strandedness: Double; CC topology: Linear; CC hypothetical: No; CC anti-sense: No; CC *source: tissue=liver; FH key Location/Qualifiers FT CDS 15..968 /product='apolipoprotein E3 precursor' FT sig_peptide 15..68 /product='apolipoprotein E3 signal peptide' FT mat_peptide 69..965 /product='apolipoprotein E3'. FEATURES source Location/Qualifiers 1..1110 /organism='Homo sapiens' /db_xref='taxon:9606' BASE COUNT 198 a 353 c 416 g 143 t ORIGIN Alignment Scores: Pred. No.: 7.33e-86 Length: 1110 Score: 1577.00 Matches: 316 Percent Similarity: 99.68% Conservative: 0 Best Local Similarity: 99.68% Mismatches: 1 Query Match: 99.24% Indels: 0 Gaps: 0 DB: 6 US-09-827-854-18 (1-317) x E00359 (1-1110)				

QY	101	THRRPVALALAGLUGLUTHRRPVALAARGLEUSERLYSGILNLEGLNALAGNALA	120
Db	315	ACCCGGGTGGCGAGAGACGCGGGACGGCTGTCCAAAGAGCTCGAGGGCGGACGGCC	374
QY	121	ATGLEUGLYALASPMEETLUALSPVALCYSGILYARGLEVALGINTYRARGLYGUVAL	140
Db	375	CGGCTGGCGGGACATGAGAGACGCTGTCCGGCCGGCTGTGTCACTAGACGGCGGACGGTG	434
QY	141	GLNALAMELEUGLYGINSERTHRTGLUGLEUARGVALARGLEUALASERTHISLEUARG	160
Db	435	CAGGCCATGCTCGGCGACAGACACCGAGAGAGCTCGGGGTGCCCTCGCCCTCCACCTGGCC	494
QY	161	LYSLEUCYSLYARGLEUENARGSPALASPASPAPLEGLNLYSARGLEUALAVALTYR	180
Db	495	AAGCTGCGTAAACGGGCTCTCCGCGATGCCATGTACCTTCAGAAACCGCTGGCACGTAC	554
QY	181	GLNALAGLYALARGGLUGLYLAGLUGLARGGLYLYLEUSERVALILEARGGLUARGLEUGLY	200
Db	555	CAGGCCGGGGGCCCCGGAGGGGCCCCGAGCGGCTTCAGCGGCCATCCGAGCGGCTGGGG	614
QY	201	PROLEUVALGLUGLNGLYARGVALARGVALARGVALAARGHVALGILSERLEUALAGLYGNPRO	220
Db	615	CCCCGTGTGGAGACAGAGGCGCGCTGGGGGGCGCCACTGTGGGCTCCCTGGCCGGCAGCGG	674
QY	221	LEUGLNGLUARGVALGLNALATRRGLYUARGLEUARGVALAARGMETSUGLUMETGLY	240
Db	675	CTACAGAGAGCGGGCCAGGCTCTGGGGCGCGCGCTGCGCGCGGATGGAGAGATGGGC	734
QY	241	SERATGTRHTRARGSPARGLEUASPRGLUVALLYSGILUGLINALAGLVALARGVALAYS	260
Db	735	AGCCGGACCCGGCAGCCGCTGGAGCAGAGGTGAAGAGACAGGTGGCGAAGTGGCGCCCAAG	794
QY	261	LEUGLUGLUGLINALAGLNGLNLARGLEUGLINALAGLUALAARGHVALAARGLEUAYS	280
Db	795	CTGGAGAGAGACAGGCCACAGATAGCGCTGCAGGCGGAGGCTTCAGAGCGCCGCTCAAG	854
QY	281	SERTTRPHEGLUARGLEUVALGLUASPMEETLNARGGLINTTRPALGLYLEUVALGULAYS	300
Db	855	AGCTGTTTCGAGACCCCTGTGGGAACATATGCAGCCGCAAGTGGGCGGGCTGTGGAGAAG	914
QY	301	VALGLNALAVALGILYTHRSERTALAPROVALPROSERASPASNHS 317	
Db	915	GTGCAGGCTGCCGTGGGCACAGCCGCCGCCCTGTGCCACGCAATATCAC 965	
RESULT 2			
LOCUS	E00823	1110 bp DNA linear	PAT 29-SEP-1997
DEFINITION			
		DNA sequence coding for human apolipoprotein E and its signal peptide.	
ACCESSION	E00823		
VERSION	E00823.1	GI:2169084	
KEYWORDS		JP 1986096997-A/1.	
SOURCE		unidentified.	
ORGANISM		unidentified.	
REFERENCE		1 (bases 1 to 1110)	
AUTHORS		Teranishi,Y., Matsui,Y., Ikeda,Y. and Kimura,M.	
TITLE		PRODUCTION OF HUMAN APOLIPOPROTEIN E-LIKE PROTEIN	
JOURNAL		Patent: JP 1986096997-A 1 15-MAY-1986;	
		MITSUBISHI CHEM IND LTD	
COMMENT		OS Human (Homo sapiens)	
		PN JP 1986096997-A/1	
		PD 15-MAY-1986	
		PF 16-OCT-1984 JP 1984216987	
		PI TERANISHI YUTAKA, MATSUI YASUSHI, IKEDA YASUOKO, KIMURA MASAKO	
		PC C12P21/00,A61K35/74,A61K37/04,C12N15/00,(C12P21/00,C12N1:19),	
		PC (C12N15/00),	
		PC (C12N1:19);	
		CC strandedness: Double;	
		CC topology: Linear;	
		CC hypothetical: No;	
		CC anti-sense: No;	
		CC *source: tissue type-Liver;	

	Key	Location/Qualifiers
FT	3'UTR	1..14
FT	sig_peptide	15..68
FT	peptide	/product='human apolipoprotein E signal'
FT	CDS	69..968
FT		/product='human apolipoprotein E'
FT		FT
FEATURES		
source	1..1110	Location/Qualifiers
	/organism="unidentified"	
	/db_xref="taxon:32644"	
BASE COUNT	198 a 353 c 416 g 143 t	
ORIGIN		
Alignment Scores:		
Pred. NO.:	7.33e-86	Length: 1110
Score:	1577.00	Matches: 316
Percent Similarity:	99.68%	Conservative: 0
Best Local Similarity:	99.68%	Mismatches: 1
Query Match:	99.24%	Indels: 0
DB:	6	Gaps: 0
US-09-827-854-18 (1-317) x E00823 (1-1110)		
QY	1 MellysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	15 ATGAAGGTCTGTGGCTGGCTTCTGTGCATCTCTGCAGAGATCCAGGCCAAGGTG	74
QY	21 GluLysAlaValAlaGluTrpGluProGluLysLeuArgLysGlnThrGluTrpGlnSer	40
Db	75 GAGCAAGCGGTGGAGACAGAGCCGAGCTGGCCAGAGACCGAGTGGCAGAGC	134
QY	41 GlyLysArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr	60
Db	135 GGCAGGCGGTGGAGACAGCGCAGCTGGCTTTGGATTAACCTGCCCTGGTGCAGACA	194
QY	61 LeuSerGluLysValGlnGluLysLeuLeuSerSerGlnValThrGlnLysLeuArgAla	80
Db	195 CTGTCTAGCAGGTGGCAGAGAGCTCTCAGCTCCAGGTCCAGCAAGTGAAGGGCG	254
QY	81 LeuMetAspGluTrpMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluLysLeu	100
Db	255 CTGATGAGCAGACCAAGAGAGATTAAGGCTTCAAAACGGAGCTGGAGAAACATG	314
QY	101 ThrProValAlaGluGluTrpArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	315 ACCCGGGTGGGAGAGAGACGGGGCAGCGCTGTCCAGAGAGCTCCAGCGGGCGAGCC	374
QY	121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal	140
Db	375 CGCGTGGGGCGGACATGGAGACGTGTGGGGCGCGCTGTGCATGACCGGGCGAGGTG	434
QY	141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	435 CAGGCATGCTCGGGCCAGACACCGAGAGACTCGGGGTCCGCTCCCTCCACCTGGCG	494
QY	161 LysLeuSlyAsrArgLeuLeuArgAspAlaAspAspLeuLysAsrArgLeuAlaValTyr	180
Db	495 AAGCTGGCTAGACGGCTCTCCCGGATGCGGATGACCTGGAGAAAGGCCCTGGAGTGTAC	554
QY	181 GlnAlaGlyAlaAsrGluGluGluAlaGluArgGlyLeuSerAlaLeuArgGluArgLeuGly	200
Db	555 CAGCGCGGGGGCGCGAGGGCGCGAGCGGGCGCTCAGCGCCATCCGCGAGCGCCCTGGGG	614
QY	201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyLysPro	220
Db	615 CCCCTGGTGGAAACAGGGCCGCTCGGGCGCGCACTGTGGGTCTCCCTGGCGGCGAGCG	674
QY	221 LeuGlnGluValArgAlaGlnAlaTyrPglyLysLeuLeuArgLysArgMetGluGluMetGly	240
Db	675 CTACAGAGAGCGGGCCAGGGCTGGGGCGAGCGGCTGGCGCGGATGGAGGAGATGGCG	734

OY	241	SerTgrrrArTAASAPRleuAsplVallysgluInVAlAglValArAlarAllys	260
Db	735	AACCGAGCCCGGAGCCTGTGGAGGTGAAGGACAGTGCCGAGGTGCGCCCAAG	794
OY	261	LeuglUGluGlInAlAgInGlnIleArgrLeuGlInAlAglAlApheGlInAlArgLeuLys	280
Db	795	CTGGAGGAGCAGGCCACAGATACGCTTCAGAGCCGAGGCTTCAGGCCCTCAAG	854
OY	281	SerrTPhegluPProLeuValGIUAspMetGlnArgrIntTPAlagLYLeuValGIuLys	300
Db	855	AGCTGGTTTCAGACCCTGTGGAAACATGACAGCCCCAGTGGCGCGGCTGTGGAGAAG	914
OY	301	ValGlnAlaValAlGlyThrsSerAlaIAProValPProserASPasnHis 317	
Db	915	GTCAGAGCTGCCGTGGCACCAAGCCGCCCTGTGCCAGCACAATCAC 965	
RESULT 3			
LOCUS	AX302545	1147 bp	DNA Linear PAT 30-NOV-2001
DEFINITION	Sequence 63 from Patent WO015177.		
ACCESSION	AX302545		
VERSION	AX302545.1	GI:17383082	
KEYWORDS	.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 Morin,P.J., Sherman-Baust,C.A., Pizer,E.S. and Hough,C.D.		
JOURNAL	Tumor markers in ovarian cancer		
	Patent: WO 0175177-A 63 11-OCT-2001;		
	THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)		
FEATURES	location/Qualifiers		
Source	1..1147		
BASE COUNT	210 a 365 c 425 g 147 t		
ORIGIN			
Alignment Scores:			
Pred. NO.:	7.59e-86	Length: 1147	
Score:	1577.00	Matches: 316	
Percent Similarity:	99.68%	Conservative: 0	
Best Local Similarity:	99.68%	Mismatches: 1	
Query Match:	99.24%	Indels: 0	
Db:	6	Gaps: 0	
US-09-827-854 -18 (1-317) x AX302545 (1-1147)			
OY	1	MettyValleuTPPrAlaAleuueValThrPhelAuAlagLyCySGlnAlalysVal	20
Db	46	ATGAAGGTTCGTGGCGCTGCTGTGTGCATTCTCGGAGGATGCGCAGGCCAAGGTG	105
OY	21	GlucInAlaValAlGIuThrgLuProGIuProGIuLeuArGrInGlnInThrgLUTrpInsEr	40
Db	106	GAGCAAGCGGTGGAGACAGACGCCGAGCCGACGTGCGCCACAGAACCGAGTGCAGAGC	165
OY	41	GlytInArgrTPGIuLeuAlaleuGLyArgrPheTrpAspyrYLEuArTrpValGIthr	60
Db	166	GGCCAGCCCTGGGAACCTGGCACTGGGTGCTTTTGGATTCCTGCCCTGGGTGCAGACA	225
OY	61	LeuSerGIuInValIGlIGluLeuLeuSerSerGlnAlATHrGlnGIuLeuArGla	80
Db	226	CTGTCTAGCAGAGGTGCAGAGAGAGTGTCTACACTCCACAGTCCACCCAGGAACCTAGGGCG	285
OY	81	LeuMetAspGIuThrMetLysGLuLeuLysAlArTryLysSerGIuLeuGIuGluInLeu	100
Db	286	CTGATGGACGAGACCATGAAGAGATTGAAGGCTTCAAAATCGAAGCTGAGAGAACACTG	345
OY	101	ThrProValAlaGIuGIuThrThrArGlarGrLeuSerLysGIuLeuGlnAlaAlaGlnAla	120
Db	346	ACCCCGGTGGGGAGGAGACCGGGCAAGGTGTCTCAAGAGAGCTGCAGCGCGCCAGGCC	405

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
Db 406 CGGCTGGGCGGCGCATGAGAGAGTGTGGCGCGCTGGTGCAATCCGGGGGAGAGTGTG 465
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 466 CAGGCCATYGTCTCGGCCAGAGACCGAGAGAGTGTGGTGCGCTCCAGCTCCGCG 525
QY 161 LysLeuGlyLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
Db 526 AAGCTGGCTAAGCGGCTCTCTCCGGATGCCGATGACCTGCAGAGAGCGCTGGCAGTAC 585
QY 181 GlnAlaGlyAlaArgGluGlyAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||
Db 586 CAGGCCGGGGCGCGAGGGGGCGGAGCGCGCTCAGCGCATCCGCGAGCGCTGGG 645
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
|||||
Db 646 CCGCTGTGTGAAACAGGGCGCGCTCGCGCCCACTGTGGCTCTCCGTGGCGCGCAGCCG 705
QY 221 LeuGlnGluArgAlaGlnAlaIleArgGlyLeuArgLeuArgAlaArgMetGluGlnMetGly 240
|||||
Db 706 CTACAGAGAGCGGCGCCAGGCTGGGGGAGCGGCTGGCGCGCGGATGAGAGATGGGC 765
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
|||||
Db 766 AGCGGAGACCGCGCGAGCGCTGGGAGAGCGGCTGGCGCGCGGATGAGAGATGGGC 825
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
|||||
Db 826 CTGAGAGAGAGAGCGCCAGCATGAGCTGTGAGGCGAGGCGCTTCCAGGCGCGCTCAAG 885
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
|||||
Db 886 AGCTGTTCAGACGCCCTGGTGGAAGACATGCAGGCGCCAGTGGCGGCTGTGTGGAAG 945
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
|||||
Db 946 GTGAGAGCTGCGCGGAGACAGCGCGCGCTGTGCCCGCAGCGCAATTCAC 996
RESULT 4
BD004278 1156 bp DNA linear PAT 31-JAN-2002
LOCUS BD004278 Apo E humanized mammal.
DEFINITION BD004278
ACCESSION BD004278.1 GI:18632239
VERSION BD004278.1
KEYWORDS JP 2001017028-A/2.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (Bases 1 to 1156)
AUTHORS Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.
TITLE Apo E humanized mammal.
JOURNAL Patent: JP 2001017028-A 2 23-JAN-2001;
MITSUBISHI CHEMICAL CORP
OS Homo sapiens (human)
PN JP 2001017028-A/2
PD 23-JAN-2001
PF 28-APR-2000 JP 2000128919
PR
PI SHINOBU FUJITA,HIROKI HAMANAKA,YUKO FUKUI,MINESUKE YOKOYAMA PC
A01K67/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10, PC
C12N15/09/C07K14/775,
PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,(C12N5/00,C12R1:91) CC
FEATURES
source FH key Location/Qualifiers
FT CDS 1..1156 Location/Qualifiers
1..1156 /organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 208 a 368 c 432 g 148 t
ORIGIN
Alignment Scores:
Pred. No.: 7 65e-86 length: 1156
Score: 1577.00 Matches: 316
Percent Similarity: 99.68% Conservative: 0
Best Local Similarity: 99.68% Mismatches: 1
Query Match: 99.24% Indels: 0
DB: Gaps: 0
US-09-827-854-18 (1-317) x BD004278 (1-1156)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
Db 61 ATGAGAGTTCTGTGGCTGGCTGGCTGTGCTGACATCTCTGGCAGAGATCCAGGCAAGTGT 120
QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
|||||
Db 121 GAGCAACCGGTGAG 180
QY 41 GlyGlnArgThrGluAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
Db 181 GGCAGCGCTGGGAG 240
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
Db 241 CTGTCTGAG 300
QY 81 LeuMetLysGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
|||||
Db 301 CTGATGAG 360
QY 101 ThrProValAlaGluGluThrThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
|||||
Db 361 ACCCGGTTGGCGAG 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
Db 421 CGGCTGGCGCGGAGCATGAG 480
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 481 CAGGCCATYGTCTCGGCCAGAGACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 161 LysLeuGlyLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
Db 541 AAGCTGTTCAGACGCCCTGGTGGAAGACATGCAGGCGCCAGTGGCGGCTGTGTGGAAG 600
QY 181 GlnAlaGlyAlaArgGluGlyAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||
Db 601 CAGGCCGGGGCGCGAGGGCGCGGAGCGCGGCTCAGCGCCATCCCGCAGCGCCCTGGG 660
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
|||||
Db 661 CCGCTGTGTGAAACAGGGCGCGCTGGGGCGGCACTGTGGCTCTCCGTGGCGCGCAGCCG 720
QY 221 LeuGlnGluArgAlaGlnAlaIleArgGlyLeuArgLeuArgAlaArgMetGluGlnMetGly 240
|||||
Db 721 CTACAGAGAGCGGCGCCAGGCTGGGGGAGAGCGGCTGGCGCGCGGATGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
|||||
Db 781 AGCGGAGACCGCGAGCGCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
|||||
Db 841 CTGAGAGAGAGCGCCAGCATGAGCTGTGAGGCGAGGCGAGGCGCTCAGGCGCTCAAG 900
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
|||||
Db 901 AGCTGTTCAGACGCCCTGGTGGAAGACATGCAGGCGCAGTGGCGGCGGCTGTGTGGAAGAG 960

QY 301 ValGlnAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 961 GTGACAGCTCCGTGGACACAGCCGCCCTGTGCTCCAGCACAATGAC 1011
 RESULT 5
 HUMAPOE3 1156 bp mRNA linear PRI 24-NOV-2000
 LOCUS Homo sapiens preapolipoprotein E (APOE) mRNA, complete cds.
 K00396
 VERSION K00396.1 GI:178850
 KEYWORDS apolipoprotein; apolipoprotein E; lipoprotein; polymorphism; very low density lipoprotein.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 355 to 1156)
 AUTHORS Breslow, J.L., McPherson, J., Nussbaum, A.L., Williams, R.W., Lojquist-Kahl, F., Karathanasis, S.K. and Zannis, V.I.
 TITLE Identification and DNA sequence of a human apolipoprotein E cDNA clone
 JOURNAL J. Biol. Chem. 257 (24), 14639-14641 (1982)
 MEDLINE 83082756
 PUBMED 6897404
 REFERENCE 2 (bases 250 to 777)
 AUTHORS Wallis, S.C., Rogne, S., Gill, L., Markham, A., Edge, M., Woods, D., Williamson, R. and Humphries, S.
 TITLE The isolation of cDNA clones for human apolipoprotein E and the detection of apoE RNA in hepatic and extra-hepatic tissues
 JOURNAL EMBO J. 2 (12), 2369-2373 (1983)
 MEDLINE 84131952
 PUBMED 6199196
 REFERENCE 3 (bases 1 to 1156)
 AUTHORS Zannis, V.I., McPherson, J., Goldberger, G., Karathanasis, S.K. and Breslow, J.L.
 TITLE Synthesis, intracellular processing, and signal peptide of human apolipoprotein E
 JOURNAL J. Biol. Chem. 259 (9), 5495-5499 (1984)
 MEDLINE 84185684
 PUBMED 6325438
 REFERENCE 4 (bases 88 to 1156)
 AUTHORS McLean, J.W., Elshourbagy, N.A., Chang, D.J., Mahley, R.W. and Taylor, J.M.
 TITLE Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing of a new variant
 JOURNAL J. Biol. Chem. 259 (10), 6498-6504 (1984)
 MEDLINE 84212473
 PUBMED 6327682
 REFERENCE 5 (bases 577 to 624)
 AUTHORS Gill, L.L., Peoples, O.P., Pearson, D.H., Robertson, F.W., Humphries, S.E., Cumming, A.M. and Hardman, N.
 TITLE Isolation and characterisation of a variant allele of the gene for human apolipoprotein E
 JOURNAL Biochem. Biophys. Res. Commun. 130 (3), 1261-1266 (1985)
 MEDLINE 85279526
 PUBMED 2992507
 REFERENCE 6 (sites)
 AUTHORS Pail, S.C. Jr., Newhouse, Y.M., Clarke, H.R., Weisgraber, K.H., McCarthy, B.J., Mahley, R.W. and Bersot, T.P.
 TITLE III hyperlipoproteinemia associated with apolipoprotein E phenotype E3/3. Structure and genetics of an apolipoprotein E3 variant
 JOURNAL J. Clin. Invest. 83 (4), 1095-1101 (1989)
 MEDLINE 89198059
 PUBMED 2539388
 COMMENT [1] corrected in [J. Biol. Chem. 258, 11422-11422 (1983)]. [J. Biol. Chem. 258, 11422-11422 (1983)] correction of [1]. [4] epsilon-3 and variant. [5] epsilon-2 allele. [6] sites: mutations resulting in type III hyperlipoproteinemia. Apo E is a component of normal human very low density lipoprotein. There are six human apo E phenotypes known to result from a single structural gene, three of the common alleles being epsilon-4,

FEATURES

source

epsilon-3 and epsilon-2. This sequence appears to be of the epsilon-3 allele. [1] argues that the apo E polymorphism involves mutations in the structural coding region: for example the epsilon-2 phenotype which is characterized by hyperlipoproteinemia is thought to result from a c to t change (arg to cys) at base 586 below [3], [5]. The sequence shown is 57% homologous with human apo A-I and 81% homologous with rat apo E. For the epsilon-4 sequence, see the separate entry.
 [J. Biol. Chem. 258, 11422-11422 (1983)] is too extensively revised relative to [1] to record the revisions in the FEATURES table. The sequence below is that of [J. Biol. Chem. 258, 11422-11422 (1983)]

Apo E is located on chromosome 19 -- Jackson, Bruns and Breslow, PNAS USA 81, 2945-2949 (1984) -- and is believed to be linked to the Apo C-II gene (see separate entry).
 The two mutations causing type III hyperlipoproteinemia (apolipoprotein E phenotype E3/3) produces substitutions of Arg for Cys at amino acid 112 and Cys for Arg at amino acid position 142. Draft entry and printed copy of sequence for [1] kindly provided by L.L. Gill, University of Aberdeen.
 Complete source information:
 Human liver [1], [J. Biol. Chem. 258, 11422-11422 (1983)], [2], [3], [4] and blood [5], cDNA to mRNA.
 Location/Organisms

gene

1..1156
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="19q13.2"
 /tissue_type="liver and blood"

CDS

61..1014
 /gene="APOE"
 /codon_start=1
 /product="preapolipoprotein E"
 /protein_id="AAB59546.1"
 /db_xref="GI:178851"
 /db_xref="GDB:600-119-691"
 /translation="MKVLMAALVTLFLGACQAKQDAVETPEPEELRQOTEMSGORM
 EIALGRFWDYLRWVOTLSEVOEELSSQVVOELRALMDETRELKAKSLELEQLP
 VAETRLRLSELQAAQARLGDMDVCGRIYQVGEVQAMIGOSTEELRYLASLRL
 KLKRLRLDADDLQRLAVYQAGAREGASIRERLGLPLVEGGRVRAATVGSILAG
 QPLQRAQAGERLARMEKSGSRTRDLDEYKEQVAVRAKLEEQAOQIRLQAAAFQ
 ARKSMPEPLVEDMKRQAGLVEKVAQVAGVSAARVPSDNIH"
 61..114
 /gene="APOE"
 /note="apolipoprotein E"
 115..1011
 /gene="APOE"
 /product="apolipoprotein E mature peptide"

sig_peptide

mat_peptide

variation

267
 /gene="APOE"
 /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3], [4]"
 /citation=[2]
 /replace="c"
 279
 /gene="APOE"
 /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3]"
 /citation=[4]
 /replace="a"

variation

variation

288
 /gene="APOE"
 /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3]"
 /citation=[4]
 /replace="a"
 369
 /gene="APOE"
 /note="in variant"
 /citation=[4]
 /replace="a"

variation

variation

409

QY	101	ThpovAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	361	ACCCGGGTGGCGGAGAGAACCGGGCAGGGCTGTCCAAAGACATCCAGCGCGCAGGCC	420
QY	121	ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal	140
Db	421	CGGCTGGCGCGGACATGAGGAGACCTGTGGCGCCCTGTGTCAATACCGCGGCGAGTGG	480
QY	141	GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValAlaGlyLeuAlaSerHisLeuArg	160
Db	481	CAGGCAAGCTCGCGCAAGACACCGAGGAGCTGGCGGTGCGCTCTCCACCTGGC	540
QY	161	LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	541	AAGCTGGCTAAGCGGCTCTCTCCGCATCCGATGACCTGCAGAAAGCGCTGGCACTGAC	600
QY	181	GlnAlaGlyAlaArgGluGlyAlaGlnArgGlyLeuSerAlaIleArgGluArgLeuGly	200
Db	601	CAGCGCGGGCGCCCGAGAGCGCGCGAGCGCGCTCACAGCCATCCGCGACCGCTTGGGG	660
QY	201	ProLeuValGluGlnGlyArgValArgAlaIleThrValGlySerLeuAlaGlyGlnPro	220
Db	661	CCCCGGGGAGACAGGGCGCGGTGGGGCGCCACCTGTGGCTCCCTGGCGGCGCAGCGG	720
QY	221	LeuGlnGluArgAlaGlnAlaIleTyrPoiYgLuArgLeuArgAlaArgMetGluGluMetGly	240
Db	721	CTACAGAGAGCGGGCCAGGCTGGGGGCGAGCGGTGCGCGCGGATGGAGAGATGGGC	780
QY	241	SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAlaArgAlaLys	260
Db	781	AGCCGGACCCCGACCGCTGTGAGAGGTGAAGGACAGGTGGCGGAGTCCGGCGCCACAG	840
QY	261	LeuGluGluGlnIleArgGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys	280
Db	841	CTGGAGGAGCGAGGCCACAGACATACGCTTGCAGGGCCGAGGCTTCCAGGGCCGCTCAAG	900
QY	281	SerTrpPheGluProLeuValGluAspMetGlnArgGlnTyrAlaGlyLeuValGluLys	300
Db	901	AGCTGGTTCGAGCCCTGTGGTGAAGACATGCAGCGCCAGTGGGCGGGCTGTGTGAGAGAG	960
QY	301	ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAspHis	317
Db	961	GTGCGAGGCTGCCGTGGGCACACGAGCGCCCTGTGCCACGACGAATTCAC	1011
RESULT 6			
LOCUS		BC003557	
DEFINITION		Homo sapiens, apolipoprotein E, clone MGC:1571 IMAGE:335712, mRNA,	
ACCESSION		BC003557	

KEYWORDS	MGC.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 1186)
REFERENCE AUTHORS TITLE	Strausberg,R. Direct Submission
JOURNAL	Submitted (20-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC) Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA
REMARK COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@cgsc.bc.ca

Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Matheson, Candice Mcleavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Seedi, Jacqueline Schein, Duane Smalus, Michael Smith, Lorraine Speede, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Zaira.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILMIL at: <http://image.llnl.gov>
Series: IRAL Plate: 6 Row: h Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4557324.

FEATURES

source

1. 1186

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="MGC:1571 IMAGE:3355712"

/tissue_type="Eye, retinoblastoma"

/clone_id="NIR_MGC_16"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

CDS

46..999

/codon_start=1

/product="apolipoprotein E"

/protein_id="AAH03557.1"

/db_xref="GI:13097699"

/translation="MKVLAALVTFLLGCAKKEQAVETPEPELRQOTMOSQRM

ELAGRFWDYLRWVQTLSEQVDELSSQVDELALMDEMKELAKKSELEQDLP

VAETFRRLSKELQAAQARLQADMEDVCGRLVQYGEVQAMIGSTELRYRLASHLR

KLRLRLDADLQRLAVYQAGAREGERSAIREGLGLVEGGRPARATVGSIG

QPLQRAQMEERLRARMEEMSRTRDLDEYKQAVARAKLEQAOQIRLQAEAFQ

ARLTSMEPELVEDLRQARAGLVKRYQAAVGSADPVPDNR"

BASE COUNT 248 a 366 c 425 g 147 t

ORIGIN

Alignment Scores:

Pred. No.: 7,86e-86 Length: 1186
Score: 1577.00 Matches: 316
Percent Similarity: 99.68% Conservative: 0
Best Local Similarity: 99.68% Mismatches: 1
Query Match: 99.24% Indels: 0
DB: 9 Gaps: 0

US-09-827-854-18 (1-317) x BC003557 (1-1186)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 46 ATGAAGGTTCTGTGGCTCGTTGCTGGACATTCCTGCGAGATGCCAGGCCAAAGTG 105
QY 21 GtGlnAlaValGtThrLupProGluLeuArgGlnGlnThrLupProGlnSer 40
DB 106 GAGCAAGCGGTGGAGACAGCGGAGCCGAGCTGCCACAGCACAGCCAGTGCGAGAC 165
QY 41 GtGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 166 GGCACGCGCTGGGAACCTGGCTCGCTTTGGGATTACCTGCGCTGGGTGCAGACA 225
QY 61 LeuSerGlnGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 226 CTGTCTGACAGTCAGAGAGAGCTGCTCCAGTCACCCAGCACTGAGAGGCG 285
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
DB 286 CTGATGAGAGACCATGAGAGAGTGAAGGCTTACAAATCGAACTGGAGAGAACATCG 345
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
DB 346 ACCCGGTGGCGAGAGACGCGGCGCTGTCCAAAGAGACTGCAGCGCGCGAGGCC 405
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140

DB 406 CGGCTGGCGCGGACATGAGAGACCTGTGCGGCCCTGTGTCAGTACCGCGGCAAGTG 465
QY 141 GlnAlaMetLeuGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 466 CAGGCGCATCTGCGGCGAGAGACCGAGAGAGCTGGCGGTGCGCTGCCCTCCACTGCGC 525
QY 161 LysLeuCysAlaArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 526 AAGCTGCGTAAGCGGCTCTCCGCGCATGTGACCTGCAGAAAGCGCTGCGAGTGC 585
QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 586 CAGGCGGCGGCGCGGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 645
QY 201 ProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
DB 646 CCCCTGTGTGAACAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 705
QY 221 LeuGlnGlnArgAlaGlnAlaTrpGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 240
DB 706 CTACAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 765
QY 241 SerArgTrpThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaLys 260
DB 766 ACCCGGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 825
QY 261 LeuGlnGlnGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
DB 826 CTGAGAGAGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 885
QY 281 SerTrpPheGlnProLeuValGlnLysAspMetGlnArgGlnTrpAlaGlnGlnGlnGlnGln 300
DB 886 ACTGCTGTGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 945
QY 301 ValGlnAlaAlaValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 317
DB 946 GTGCAAGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 996

RESULT 7
LOCUS E08423 1110 bp RNA linear PAT 29-SEP-1997
DEFINITION DNA coding human apolipoprotein E3.
E08423
ACCESSION E08423.1 GI:2176540
VERSION JP 1994315392-A/1.
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
Moriimoto, H. and Teranishi, Y.
METHOD FOR PRODUCING APOLIPOPROTEIN
Patent: JP 1994315392-A 1 15-NOV-1994;
MITSUBISHI KASEI CORP
OS Homo sapiens (human)
PN JP 1994315392-A/1
PD 15-NOV-1994
PF 11-JUN-1985 JP 1994015433
PI MORIMOTO HIRONORI, TERANISHI YUTAKA
PC C12P21/02, C07K13/00, C12N5/10, C12N15/12, C12N15/18, (C12P21/02,
PC C12R1:91),
PC (C12N5/10, C12R1:91);
CC strandedness: Double;
CC topology: Linear;
FH key
FH Location/Qualifiers
FT source 1..1110
FT sig_peptide 15..68
FT mat_peptide 69..968
FT /product="human apolipoprotein E3".
FEATURES
source 1..1110
Location/Qualifiers

		/organism="Homo sapiens" /db_xref="taxon:9606"			
BASE COUNT	198 a	355 c	414 g	143 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	1,91e-85	Length:	1110		
Score:	1570.00	Matches:	315		
Percent Similarity:	99.37%	Conservative:	0		
Best Local Similarity:	99.37%	Mismatches:	2		
Query Match:	98.80%	Indels:	0		
DB:	6	Gaps:	0		
US-09-827-854-18 (1-317) x E08423 (1-1110)					
Oy	1	MettYValLeuTrrPALAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20		
Db	15	ATGAAGGTTCTGTGGGCTGCTTGGTCTGCTACATTCTCGCAGAGATGCCAGGCCAAGGTG	74		
Oy	21	GlnGlnAlaValaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrrPglInser	40		
Db	75	GAGCAAGGGGTGGACACAGAGCGGAGCCGACGTGGCCACAGACCGAATGGCAGAGC	133		
Oy	41	GlyGlnAtrgTrrPgluLeuAlaLeuGlyLyrPheTrrPaspTrrLeuAtrgTrrValGlnThr	60		
Db	135	GGCCAGCCCTGGGAACCTGGCACTGGGTGCTTTGGCTATCCATGCGCTGGTGCACAGCA	194		
Oy	61	LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValTrrGlnGluLeuArgAla	80		
Db	195	CTGTGTGACAGGTGCAGAGAGAGCTGTCTCACTGCCAGGTCACCGAAGACTAGGGGCG	254		
Oy	81	LeuMetAspGluTrrMetLysGluLeuLysAlaTrrLysSerGluLeuGluGluGlnLeu	100		
Db	255	CTGATGGACGAGACCATAGAGAGTTGAAGGCTTACAATCGGAACCTGGAGAACACTG	314		
Oy	101	ThrProValAlaGluGluThrArgAlaAtrgLeuSerLysGluLeuGlnAlaGlnAla	120		
Db	315	ACCCGCGTGGCGAGAGAACCGGGCACGGCTGTCTCAAGAGAGCTGCAGAGCGCGCACGGCC	374		
Oy	121	ArgLeuGluLysAlaAspMetGluAspValLysGlyLyrArgLeuValGlnTrrArgGlyVal	140		
Db	375	CGGGTGGCGCGGACATGAGAGAGCTGTGGCGCGCCCTGGGTGCATGCCCGGGGAGGGTG	434		
Oy	141	GlnAlaMetLeuGlyGlnInserThrGluGluLeuArgValAtrgLeuAlaSerHisLeuArg	160		
Db	435	CAGGCCATGCTCGGCCAAGACACCGAGAGAGCTGGGGGTGGCCCTCGCCTCCACACTGCC	494		
Oy	161	LysLeuCysLysAtrgLeuLeuArgAspAlaAspAspLeuGlnLysAtrgLeuAlaValTrr	180		
Db	495	AAGCTGCTAAGCGGCTCTCCCGGATGCCCCATGACCTGCAGAGAGCGCTGGCAGTCTAC	554		
Oy	181	GlnAlaGlyAlaAtrgGluGlyAlaGluArgGlyLeuSerAlaIleAtrgGluArgLeuGly	200		
Db	555	CAGCGGGGGCGCGGAGAGGGCGCGAGCGCGCTTCAGCGCATCCCGAGCGGCTGGGG	614		
Oy	201	ProLeuValGlnGlnGlyArgValArgAlaIleAtrhValGlySerLeuAlaGlyGlnPro	220		
Db	615	CCCCGTGGTGGAAACAGGGCGCGGTGCGGGCCGCCACTGTGGGTCTCTTGCCCGGCCAAGC	674		
Oy	221	LeuGlnGluAtrgAlaGlnAlaIleAtrPglLysLyrArgLeuArgAlaAtrMetGluGluMetGly	240		
Db	675	CTACAGAGCGGGGCCACAGGCTGGGGGGAGGGCGCTGGCGCGCGGATGGAGAGATGGGC	734		
Oy	241	SerAtrgTrrAtrgAspAtrgLeuAspGluValLysGluGlnValAlaGluValAtrgAlaLys	260		
Db	735	AGCCGGACCCCGACCGCTGTGAGAGGTGAAGAGACAGGTGGCGGAGGTGGCGGCACAG	794		
Oy	261	LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaAtrgLeuLys	280		
Db	795	CTGGAGAGACAGGCCCCAGCAGATACGCTGTGAGGCCCAAGGCTTCCAGGCCCGGCTCAAG	854		
Oy	281	SerTrrPheGluProLeuValGluAspMetGlnAtrgGlnTrrPAlaGlyLeuValGluLys	300		

Db	855	ACCGTGGTTCAGACCCCTGGTGGAGACATGCAGCCCACTGGGCGGCGTGGTGGAGANG	914
Oy	301	ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis	317
Db	915	GTGCAAGCGTGGCGTGGGACACGAGCGCGCGCTGTGGCCAGGACAAATCAC	965
RESULT 8	BD004277		
LOCUS	BD004277	1156 bp	linear
DEFINITION	Apo E humanized mammal.		
ACCESSION	BD004277		
VERSION	BD004277.1	GI:18632238	
KEYWORDS	JP 2001017028-A/1.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 1156)		
JOURNAL	Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.		
COMMENT	Apo E humanized mammal		
	Patent: JP 2001017028-A 1 23-JAN-2001;		
	MITSUBISHI CHEMICAL CORP		
	OS Homo sapiens (human)		
	PN JP 2001017028-A/1		
	PD 23-JAN-2001		
	PF 28-APR-2000 JP 2000128919		
	PR		
	PI SHINOBU FUJITA,HIROKI HAMANAKA,YUKO FUKUI,MINESUKE YOKOYAMA		
	PC A01K67/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10,PC		
	C12N15/09//C07K14/775,		
	PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,(C12N5/00,C12R1:91) CC		
FEATURES	FT	Key	Location/Qualifiers
	source	1..1156	Location/Qualifiers
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
BASE COUNT	208 a	367 c	432 g
ORIGIN			149 t
Alignment Scores:			
Pred. No.:	2,29e-85	Length:	1156
Score:	1569.00	Matches:	315
Percent Similarity:	99.37%	Conservative:	0
Best Local Similarity:	99.37%	Mismatches:	2
Query Match:	98.74%	Indels:	0
DB:	6	Gaps:	0
US-09-827-854-18 (1-317) x	BD004277 (1-1156)		
Oy	1	MellysValleuTTPAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal	20
Db	61	ATGAAGGTTCTGTGGCGCTGCTGTGCATTCCTGGAGAGATGCGACCAAGTG	120
Oy	21	GlunAlaValAlaGlyThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer	40
Db	121	GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCGCCACAGACAGCCAGTGGCAGAGC	180
Oy	41	GlyGlnATGTPGluLeuAlaLeuGlyArgPheThrAspTyrLeuArgTyrValGlnThr	60
Db	181	GGCCAGCGCTGGGAAGTGGCACTGGGTGCGTTTGGATTTACCTGCGCTGGGTGCAGACA	240
Oy	61	LeuSerGluGlnValGlnGluGluLeuSerSerGlnValThrGlnGluLeuArgAla	80
Db	241	CTGTCTGAGCGAGTGCAGAGAGAGTGTCTCAGCTCCAGTCCACCCAGAACTAGAGGCG	300
Oy	81	LeuMetAspGluThrMetTyrGluLeuLysAlaTyrTyrLysSerGluLeuGluGlnGluLeu	100
Db	301	CTGATGGACGAGACCATTAAGAGAGTTGAAGGCTTCAATCGAATCTGAGAGCAACTG	360
Oy	101	ThrProValAlaGluGluThrArgAlaArgLeuSerTyrGluLeuGlnAlaGlnAla	120

Db 361 ACCCGGTGGCGAGGAGACGCGGCGACGGCTGCCAAGAGAGCTGCAGCGCGGCGAGCC 420
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnIrrArgGlyVal 140
Db 421 CGGCTGGCGCGACATGAGGAGCGGTGGCGCGCTGGTGCAGTACCAGCGCGGAGGTG 480
Qy 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCGATGCTGGCGGACAGACACCGAGAGAGCTGGGGTGGCTCCCTCCACCTGGCG 540
Qy 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AACCTGCGTAAAGCGGCTCCCTCCGATGCCGATGACCTGCAGAAAGTCCGCGAGGTAC 600
Qy 181 GlnAlaGlyAlaArgGluGluValArgGlyLeuSerAlaLeuArgGluValArgLeuGly 200
Db 601 CAGGCGGGGCGCGGAGGCGCGGAGCGGCTCAGCGCCATCCGCGCGCGCGCGG 660
Qy 201 ProLeuValGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGTGTAACAGGCGCGCGCTGGGCGCGCCACTGTGGGCTCCCTGGCGCGCGAGCCG 720
Qy 221 LeuGlnGluArgAlaGlnAlaTrpGlyArgLeuArgAlaArgMetGluGluMetGly 240
Db 721 CTACAGAGGCGGCGCGAGGCTGGGCGGAGCGGCTGGCGCGCGATGAGAGATGGCG 780
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAlaArgAlaLys 260
Db 781 AGCGGAGCGCGGACCGCTGGAGGAGGTGAAGAGAGAGAGTGGGAGGCGCGCGCAAG 840
Qy 261 LeuGlnGluGlnAlaGlnGlnAlaArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 841 CTGGAGAGAGGCGCGAGAGATAGCCCTGSCAGAGCGGAGGCTCCAGCGCGCGCTCAG 900
Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 901 AGCTGTTGACGCGCTGGGAGAGACATGACAGCGCGAGGCGCGGCTGGTGGAGAG 960
Qy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTCCAGGCTGCGCGTGGGACACAGCGCGCGCTGTGCCCGCAGCAATCATC 1011

RESULT 9 AX333278 1157 bp DNA linear PAT 09-JUN-2002
LOCUS AX333278 Sequence 3787 from Patent WO0194629.
DEFINITION AX333278
ACCESSION AX333278 GI:18123912
VERSION AX333278.1 GI:18123912
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horigan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 3787 13-DEC-2001;
FEATURES
source location/Qualifiers
1. 1157
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN

Alignment Scores:
Pred. No.: 2,63e-85
Score: 1568.00
Percent Similarity: 99.058
Best Local Similarity: 99.058
Query Match: 98.688

Length: 1157
Matches: 314
Conservative: 0
Mismatches: 3
Indels: 0

DB: 6 Gaps: 0
US-09-827-854-18 (1-317) x AX333278 (1-1157)
Qy 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAGTTCTGTGGGCTGGCTGGTGTGTCACATTCCTGGACAGAGTCCAGGCGCAAGGTG 121
Qy 21 GlnGlnAlaValGlnThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 122 GAGCAAGCGGTGGAGACAGAGCGGAGCGCGGAGCTCCGCGCAGAGAGCCGAGTGGCAGAG 181
Qy 41 GlnArgTrpGluLeuAlaLeuGlnArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 182 GGCACGCTGGGAAGTGGCACTGGGTCCGTTTGGGATTAACCTGGCTGGGTGGAGACA 241
Qy 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db 242 CTGTGACAGAGTGCAGAGAGAGTGTACAGCTCCCAAGTCAACCAAGAACTGAGGCGG 301
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
Db 302 CTGATGACAGAGCACTGAAGGATTGAAGGCTTACAAATCGGAACTGAGAGAACTG 361
Qy 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
Db 362 ACCCGGTGACGAGAGAGAGCGCGGCGACGGCTGTCCAAAGAGCTGCAGAGCGCGAGGCC 421
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnIrrArgGlyVal 140
Db 422 CGGCTGGCGCGGACATGAGAGAGCGTGGCGGCGCGCTGGTGCAGTACCAGCGCGAGGTG 481
Qy 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCGCATCTCGCGCGAGAGCACCGAGGAGCTCGGGTGGCTCCCTCCCACTGGCGC 541
Qy 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AACCTGCGTAAAGCGGCTCCCTCCGATGACCTGCAGAGAGCGCTGGCAGTGTAC 601
Qy 181 GlnAlaGlyAlaArgGlnGlnAlaArgGlyLeuSerAlaLeuArgGluArgLeuGly 200
Db 602 CAGGCGGGGCGCGGAGGCGCGGAGCGCGGCTCAGCGCATCCGCGAGCGCGCTGGGG 661
Qy 201 ProLeuValGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 662 CCCCTGTGTAACAGGCGCGCGCTGGGCGCGCCACTGTGGGCTCCCTGGCGCGCGAGCCG 721
Qy 221 LeuGlnGluArgAlaGlnAlaTrpGlyArgLeuArgAlaArgMetGluGluMetGly 240
Db 722 CTACAGAGAGGCGCGGACGCTGGGCGAGCGGCTCCGCGCGATGGAGAGATGGGC 781
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValAlaArgAlaLys 260
Db 782 AGTCGAGCGCGGACCGCTGGAGAGGTGAAGAGACAGTGGCGGAGGTGGCGCGCAAG 841
Qy 261 LeuGlnGlnGlnAlaGlnGlnAlaArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 842 CTGGAGAGACAGCGCGAGAGATAGCGCTGAGAGCGCGAGGCTCCAGGCGCGCTCAAG 901
Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 902 AGCTGTTGAGCGCGCTGGTGAAGACATGCAAGCGCAATGGCGCGGCTGGTGGAGAG 961
Qy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTCCAGGCTGCGGTGGGACACAGCGCGCGCTGTGCCCGCAGGACATCATC 1012

RESULT 10 AX409597 1157 bp DNA linear PAT 14-JUN-2002
LOCUS AX409597 Sequence 2244 from Patent WO0229103.
DEFINITION AX409597
ACCESSION AX409597

VERSION AX409597.1 GI:21442302
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Alvarres,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2244 11-Apr-2002;
GENE LOGIC INC (US)
FEATURES
source location/Qualifiers
1..1157
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="EMBL/Genbank Accession No. M12529"
BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN
Alignment Scores:
Pred. No.: 2.63e-85 Length: 1157
Score: 1568.00 Matches: 314
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 99.05% Mismatches: 3
Query Match: 98.68% Indels: 0
DB: Gaps: 0
US-09-827-854-18 (1-317) x AX409597 (1-1157)
QY 1 MettlysalvleutrpalaalaleuleuValThrpheulaaglycysglnalalysVal 20
Db ATGAAGTTCTGTGGCTGGCTGGCTGCACATCTCTGCAGAGATCCAGGCCAAGGTG 121
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnInThrGluTrpGlnSer 40
Db 122 GAGCAAGCGGTGGAGAGAGAGCCGAGCCGAGCTGCGCCAGCAGACCGAGTGGCAAGC 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 182 GGCACAGCGTGGCAACGTCAGTGGCTGTTGGATTACCTGGCTGGTGGCAACA 241
QY 61 LeuSerGluGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db 242 CTGTCTGAGCAGGTGGAGAGAGCTGTCTCAAGTCCCAAGTCAACCAAGAGAGAGCG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
Db 302 CTGATGGACAGACCATGAGAGAGTTGAAGCCCTACAAATCGGAACCTGGAGAACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 362 ACCCGGTACCGAGGAGAGCGGGGCGAGCTGTCCAAAGAGCTGCAGAGCGCGCAGGCG 421
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlnVal 140
Db 422 CGGCTGGGCGCGGACATGAGAGAGTGTGGCGCGCTGTGCACTACCCGCGGAGAGTG 481
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCCATGCTCGGCCAGACACCGAGAGCTGGGGTGGCGCTCGCTCCCACTCGCGC 541
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AAGCTGTGAAGCGGCTCCCTCCGATCCCGATGACCTGCAGAAAGCCCTGGCAGTGTAC 601
QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
Db 602 CAGCGCGGCGCGCGAGAGGCGCGAGCGCGCGCTCAAGCCATCCGCGAGCGCGTGGGG 661
QY 201 ProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
Db 662 CCCTGTGTGAACAGGCGCGCTGGCGCGCGCTGTGGCTGTCCCTGGCGCGCGCAGCGG 721

QY 221 LeuGlnGlnAlaArgAlaGlnAlaTTPGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
Db 722 CTACAGAGAGGCGCGCGAGCGCTGGCGGAGCGCTGTGGCGCGGAGAGAGATGGCG 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
Db 782 AGTCGGACCGCGCGAGCGCTGGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 841
QY 261 LeuGlnGlnGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
Db 842 CTGGAG 901
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 902 AGCTGTGTGAG 961
QY 301 ValGlnAlaAlaValGlnThrSerAlaAlaProValProSerAspHis 317
Db 962 GTGCAGGCTGCGTGGGACACAGCGCGCGCTGTGGCCAGCCGCAATCTAC 1012
RESULT 11
LOCUS 115975 1157 bp DNA linear PAT 03-APR-1996
DEFINITION Sequence 1 from patent US 5472858.
ACCESSION 115975
VERSION 115975.1 GI:1250883
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1157)
AUTHORS Attie,A.D., Gretch,D.G., Sturley,S.L. and Beckage,N.E.
TITLE Production of recombinant proteins in insect larvae
JOURNAL Patent: US 5472858-A 1 05-DEC-1995;
FEATURES
source location/Qualifiers
1..1157
/organism="unknown"
BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN
Alignment Scores:
Pred. No.: 2.63e-85 Length: 1157
Score: 1568.00 Matches: 314
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 99.05% Mismatches: 3
Query Match: 98.68% Indels: 0
DB: Gaps: 0
US-09-827-854-18 (1-317) x 115975 (1-1157)
QY 1 MettlysalvleutrpalaalaleuleuValThrpheulaaglycysglnalalysVal 20
Db 62 ATGAAGTTCTGTGGCTGGCTGGCTGCACATCTCTGCAGAGATCCAGGCCAAGGTG 121
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnInThrGluTrpGlnSer 40
Db 122 GAGCAAGCGGTGGAGAGAGAGCCGAGCCGAGCTGCGCCAGCAGACCGAGTGGCAAGC 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 182 GGCACAGCGTGGCAACGTCAGTGGCTGTTGGATTACCTGGCTGGTGGCAACA 241
QY 61 LeuSerGluGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
Db 242 CTGTCTGAGCAGGTGGAGAGAGAGTGTCTCAAGTCCCAAGTCAACCAAGAGAGAGCG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
Db 302 CTGATGGACAGACCATGAGAGAGTTGAAGCCCTACAAATCGGAACCTGGAGAACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 362 ACCCGGTACCGAGGAGAGCGGGGCGAGCTGTCCAAAGAGCTGCAGAGCGCGCAGGCG 421

QY 121 ATGLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 |||||
 Db 422 CGGCTGGGGCGGACATGAGGAGCGGTGGGGCCGCTGGTGCAGTACCGCGGAGGTG 481
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||
 Db 482 CAGGGCATCTCGGGCAGAGCACCGAGGAGCTGGGGTCTGGCTCCCTCCACCGTGGCG 541
 QY 161 LysLeuCysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||||
 Db 542 AAGCTGCGTAAGCGGCTCTCGCGATCCCGATGACCTCAGAAAGCGCTGGCAGGTAC 601
 QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 |||||
 Db 602 CAGCGCGGGGGCGCGAGGGCGCGGCGGCTCAGCGCCATCCGCAAGCGCTGGGG 661
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 |||||
 Db 662 CCCCTGGTGAACAGGGCGCGGCGGCGGCTGTGGCTCCCTGGCGCGCGCG 721
 QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyArgLeuArgAlaArgMetGluGlnMetGly 240
 |||||
 Db 722 CTACAGAGCGGGCCAGGCTGGGGCGAGCGGCTGCGCGGATGAGAGATGGGC 781
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 |||||
 Db 782 AGTGGAGCCCGGAGCGCGCTGGAGCGAGTGAAGAGACAGTGGCGAGGTGGCGCGCAAG 841
 QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 |||||
 Db 842 CTGGAGAGACAGGCGCGAGATACGCTTCGAGCGCGAGCGCTTCAGGCGCGCGCTCAAG 901
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 |||||
 Db 902 AGCTGTTGAGCCCTGGTGGAGCATGACAGCGCACTGGGCGGGCTGGTGGAGAG 961
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 |||||
 Db 962 GGGCAGGCGTGGCGGCGACAGCGCGCGCTGTGCCAGCAGACATATAC 1012

RESULT 12
 HUMAPOE 1157 bp mRNA linear PRI 08-AUG-1995
 LOCUS Human apolipoprotein E mRNA, complete cds.
 DEFINITION M12529
 ACCESSION M12529.1 GI:178848
 VERSION
 KEYWORDS apolipoprotein.
 SOURCE Homo sapiens (clone: pHAEl112,178,813). male 57-year old liver
 CDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1157)
 AUTHORS McLean,J.W., Elshourbagy,N.A., Chang,D.J., Mahley,R.W. and Taylor,J.M.
 TITLE Human apolipoprotein E mRNA. CDNA cloning and nucleotide sequencing
 of a new variant
 JOURNAL J. Biol. Chem. 259 (10), 6498-6504 (1984)
 MEDLINE 84212473
 PUBMED 6327682
 FEATURES
 source location/Qualifiers
 1..1157
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="19q13.2"
 /clone="pHAEl112,178,813".
 /sex="male"
 /tissue.type="liver"
 /dev_stage="57-year old"
 1..1157
 /gene="APOE"
 <1..1157

gene
 mRNA

CDS
 /gene="APOE"
 /note="apoe mRNA; G00-119-691"
 62..1015
 /gene="APOE"
 /note="precursor"
 /codon_start=1
 /product="apolipoprotein E"
 /protein_id="AAB59518.1"
 /db_xref="GI:178849"
 /db_xref="GDB:G00-119-691"
 /translation="MKVLMALLVTFLLGCAKKYBOAVTEPPEPLRQOTEMSGORM
 ELATGRFWDYLRWQTFLEQVQEBLLSSQVQELRALMDETMKLAKVSELEQLTP
 VAETFRRLSELEQARLQADMEQVGRVQYRGVQAMLGSTEEELRYRLSLR
 KLRLRLDPDLQRLRLVQAGAREGERSLIREPLGLVEQGRVRAVTSGLG
 QPLERAQAMGERLRARMEGSRTRDLQVKEQVAVRAKLEBQAOQIRLQAEAFQ
 ARLSWEPEPLVEDMQRQNAGLVEKYQAIVGTSADVPSPDNH"
 62..115
 /gene="APOE"
 /note="G00-119-691"
 116..1012
 /gene="APOE"
 /product="apolipoprotein E"
 /note="G00-119-691"
 370
 /gene="APOE"
 /note="a in pHAEl112,1781; g in pHAEB13; G00-119-691"
 /replace="g"
 410
 /gene="APOE"
 /note="a in pHAEl112,1781; g in pHAEB13; G00-119-691"
 /replace="g"
 569
 /gene="APOE"
 /note="c in pHAEl112,1781; g in pHAEB13; G00-119-691"
 /replace="g"
 784
 /gene="APOE"
 /note="c in pHAEl112,1781; c in pHAEB13; G00-119-691"
 /replace="c"
 859
 /gene="APOE"
 /note="g in pHAEl112,1781; a in pHAEB13; G00-119-691"
 /replace="a"
 variation
 BASE COUNT 212 a 370 c 426 g 149 t
 ORIGIN 30 bp upstream of BstNI site.
 Alignment Scores:
 Pred. No.: 2.63e-85 Length: 1157
 Score: 1568.00 Matches: 314
 Percent Similarity: 99.05% Conservative: 0
 Best Local Similarity: 99.05% Mismatches: 3
 Query Match: 98.68% Indels: 0
 DB: 9 Gaps: 0
 US-09-827-854-18 (1-317) x HUMAPOE (1-1157)
 QY 1 MetLysValLeuTrpAlaAlaLeuLeuValIrrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||||
 Db 62 ATGAAGGTTCTTGGGGCTGGCTGTGTCACATTCCTGGCAGATGCCAGGCAAGGTG 121
 QY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 |||||
 Db 122 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTGCGCAGACAGACCGAGTGGCAGAC 181
 QY 41 GlyIleArgTrpGluLeuAlaLeuGlnIleArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 |||||
 Db 182 GCCACAGCGCTGGAGACTGCACACTGGCTTTGGAGTTACTCGCTGGCTGGCAGACA 241
 QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 |||||
 Db 242 CTGTGTGAGCAGGTGAGAGAGCTGCTCCACAGTCAACCAAGTCAACCAAGTCAAGGCG 301
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100

|||||
Db 302 CTGATGACAGACCAAGAGGAGTTGAGCCCTACAAATCGGAACAGGAGCAACTG 361
Qy 101 TheProvalAlaGluGluThrArgAlaArgLeuSerLysGluGluGlnAlaGlnAla 120
Db 362 ACCCGGTACGAGGAGGAGACCGGGCACGGCTGTCCAAAGAGCTGCAGACGGCCAGGCC 421
Qy 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 422 CGGCTGGGCGCGACATGAGAGACGTGTGGCGCGCTGTGTGACAGTACCGCGGAGAGTG 481
Qy 141 GlnAlaMetLeuGlnGlnSerThrGluGluGlnLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCCATGCTCGGCCAGACACCGAGAGACGTGGGTGGCGCTCGCTCCACCTGCGC 541
Qy 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AACCTGGCTAAGCGGCTCTCCCGCATCCGATGACCTGCAGAAAGCCCTGGCAGTAC 601
Qy 181 GlnAlaGlnAlaArgGlnGlnAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 602 CAGGCCGCGGCCGCGAGGCGCGGACGCGGCTCAGGCCCATCCGCGAGCGCGCTGGG 661
Qy 201 ProLeuValGluGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 662 CCCCTGGTGAACAGGCGCCGCTGGCGGCGCCACATGTGGCTCCCTGGCGCGGACCGC 721
Qy 221 LeuGlnGluArgAlaGlnAlaArgGlyGlyLeuArgLeuArgAlaArgMetGluGluMetGly 240
Db 722 CTACAGAGAGCGGCGCCAGGCGCTGGGGGAGGAGCGGCTGGCGCGGATGGAGAGATGGCG 781
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 782 AGTCGGAGCCCGCAGCCCTCGAGCAGAGTGAGAGCAGTGCGGAGTGCGCGGCCAAG 841
Qy 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaGlnAlaArgLeuLys 280
Db 842 CTGAGAGAGAGGCCAGCAGATACGCTGACAGGCCAGGCTTCACAGGCCCTCAAG 901
Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 902 AGCTGGTTCAGGCCCTCGGTGGAGAACATGCACAGCGCAGTGGCGCGGTGGTGGAGAG 961
Qy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGACAGCTGCGCTGGGCGACACAGCGCGCCCTGTGGCCAGCAGCAATCAC 1012
RESULT 13
BD004279 1156 bp DNA linear PAT 31-JAN-2002
LOCUS Apo E humanized mammal.
DEFINITION BD004279
ACCESSION BD004279.1 GI:18632240
VERSION JP 2001017028-A 3
KEYWORDS JP 2001017028-A/3.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1156)
Apo E humanized mammal
Patent: JP 2001017028-A 3 23-JAN-2001;
MITSUBISHI CHEMICAL CORP
OS Homo sapiens (human)
PN JP 2001017028-A/3
PD 23-JAN-2001
PE 28-APR-2000 JP 2000128919
PR
PI SHINDBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA PC
AO1k67/027.A61k45/00.A61p25/28.A61p43/00.C12N5/10, C12N5/09//C07K14/775,
C12N15/09//C07K14/775,
PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,(C12N5/00,C12R1:91) CC

.. Alignment Scores:
Pred. No.: 3,97e-85 Length: 1156
Score: 1565.00 Matches: 315
Percent Similarity: 99.378 Conservative: 0
Best Local Similarity: 99.378 Mismatches: 2
Query Match: 98.49% Indels: 0
DB: Gaps: 0
US-09-827-854-18 (1-317) x BD004279 (1-1156)
Qy 1 MetLysValLeuThrProAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAGAGTCTGTGGGTGGCTGTGTGTGTCATCTCTGGCAGAGATGCCAGGCCAAGTTC 120
Qy 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrGlnThr 40
Db 121 GAGCAAGCGGTGAGACAGACAGCGGAGCCGACCTGCGCCAGACAGACGAGTGGCAGAGC 180
Qy 41 GlyLAlaTrpPheGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 181 GGCACAGCGCTGGAGACTGCGCAGCTGGCTTTGGGATTACTGCGCTGGGTGCAGACA 240
Qy 61 LeuSerGluGlnValGlnGlnLeuLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db 241 CTGTCTGAGCAGAGTGCAGAGAGAGTGTCTGACCTCCAGATCCAGGAACTGAGAGCGG 300
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
Db 301 CTGATGAGCAGACCAATGAAGGAGTTGAAGCCTTACAAATCGGAATGGAGGAACAATG 360
Qy 101 TheProValAlaGluGlnThrArgAlaArgLeuSerLysGluGlnAlaAlaGlnAla 120
Db 361 ACCCGGTGGCGGAGGAGAGCGCGGCGACGGCTGTCCAAAGAGCTGCAGAGCGCGCAGGCC 420
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 421 CGGCTGGGCGCGACATGAGGAGCTGCGCGGCGCCCTGTGTGACATGACCGCGCGAGTG 480
Qy 141 GlnAlaMetLeuGlnGlnSerThrGluGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTCGGCCAGAGACACCGAGAGACTGGGGTGGCTCGCTCCACTGCGC 540
Qy 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AACCTGGCTAAGCGGCTCTCCCGCATCCGATGACCTGCAGAAAGCGCTGGCAGATGTAC 600
Qy 181 GlnAlaGlyAlaArgGlnGlnAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 601 CAGCGCGGCGCGCGAGGCGCGGCGGCTGACGCGCTCAGCGCCATCCGAGAGCGCTGGGG 660
Qy 201 ProLeuValGlnGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
Db 661 CCCCTGGTGAACAGGCGCGCGCTGGCGGCGCCACTGTGGCTCCCTGGCGCGCGAGCGG 720
Qy 221 LeuGlnGluArgAlaGlnAlaArgGlyGlyLeuArgLeuArgAlaArgMetGluGluMetGly 240
Db 721 CTACAGAGAGCGGCGCCAGGCGCTGGGGGAGGAGCGGCTGGCGGATGGAGAGATGGCG 780
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
Db 781 AGCGGAGACCGCGAGCGCTGTGACGAGTGAAGAGACAGTGGCGGAGCTGGCGCGCCCAAG 840
Qy 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
|||||

Db 841 CTGGAGGACAGCCAGCATACGCTGACAGCCGCTTCCAGCCCGCCTCAAG 900
Oy 281 serTTPhegluProleuValgluAspMetgluArggluTrrPalagluLeuValgluIlys 300
Db 901 AGCTGTTGACCCCTGGTGGAGACATGCACGCCCATGTGGCCGCTGGTGGAGAAG 960
Oy 301 ValgluAlaValgluIlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCGGTGGACACAGCGCCGCTGTGCCAGCAGCATATCAC 1011
RESULT 14
MFAPOE 1178 bp mRNA linear PRI 31-MAR-1995
DEFINITION Monkey mRNA for apolipoprotein E.
ACCESSION X13887
VERSION X13887.1 GI:38054
KEYWORDS apolipoprotein; apolipoprotein E.
SOURCE Macaca fascicularis.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
Macaca.
REFERENCE 1 (bases 1 to 1178)
AUTHORS Marotti,K.R.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1988) Marotti K.R., The Upjohn Co, 7242 209 713,
301 Henrietta Street, Kalamazoo, MI 49008
2 (bases 1 to 1178)
Marotti,K.R., Whitted,B.E., Castle,C.K., Polites,H.G. and
Melchior,G.W.
TITLE Nucleotide sequence of the cynomolgus monkey apolipoprotein E cDNA
JOURNAL Nucleic Acids Res. 17 (4), 1778 (1989)
MEDLINE 89160349
PUBMED 2922300
COMMENT Data kindly reviewed (20-Mar-1989) by Marotti K.R.
FEATURES
source
Location/Qualifiers
1..1178
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="1"
/tissue_type="liver"
/clone_idb="lambda gtcl1"
83..1036
/codon_start=1
/product="proapolipoprotein E (AA -18 to 299)"
/protein_id="CAA32092.1"
/db_xref="GI:38054"
/db_xref="SWISS-PROT:P10517"
/translation="MKVLMALVTFILGCOAKVEQVPEPETEPELRQAEGSGQPM
ELALGRWDYLRWVQLSEVOEELISPOVQELSTLMDETMKELKAKSELEBOLSP
VAETRRARLSKELQAAOARLGADMEDVSRLOYRSEVOAMLGOSTEELRARRASHLR
QLKRRLRDADDLQRLAVQAGAREGAREVSRVLEGLVLDGRRARRATVSSLAS
QPLDERAOLAEERLRARMEKGRSRDRDLDEKEDVAEVRKLEQAOQISLQNEAFO
ARKLSWEPEPLVEDMOKWAGLVEKVOAAGVASTAVPIDNH"
sig_peptide
mat_peptide
misc_feature
BASE COUNT 217 a 373 c 426 g 162 t
ORIGIN

Alignment Scores:
Pred. No.: 7.49e-78 Length: 1178
Score: 1443.00 Matches: 295
Percent Similarity: 93.69% Conservative: 2
Best Local Similarity: 93.06% Mismatches: 20
Query Match: 90.81% Indels: 0
DB: 9 Gaps: 0
US-09-827-854-18 (1-317) x MFAPOE (1-1178)
Oy 1 MetlvsValleuTrrPalalaLeuValIlyThrPheleuAlagluCysgluAlaIlysVal 20

Db 83 ATGAAGGTTCTGTGGGCTCGTTGGCTGTCATTCCTTGGCAGAGATCCAGGCGCAAGCTG 142
Oy 21 GluAlaValAlaValThrIleuProgluIleuArggluIleuThrIleuPrgInsar 40
Db 143 GAGCAACCGGTGAGCCAGACAGAGACCGGAGCTTGCACAGAGCTAGAGGCGGAGAC 202
Oy 41 GlyAluArgTrrPgluLeuAlaLeuGluArgPheTrrAspTrrLeuArgTrrValGlnThr 60
Db 203 GGCCAGCCCTGGGAGGTGGACACTGGGTCTTTTGGGATTACCTGGCGGTGGAGACA 262
Oy 61 LeuSerGluGluValGluGluGluLeuLeuSerSerGluValThrGluGluLeuArgAla 80
Db 263 CTGTCTGACAGAGTGCAGAGAGAGCTGTACAGCCCGCCAGCATCCAGGAACTGACAGACG 322
Oy 81 LeuMetAspGluThrMetIysGluLeuIysAlaIlyThrIysSerGluLeuGluGluGlu 100
Db 323 CTGATGGAGAGACCATGTAAGAGGTTGAAGGCTTACAAATCGGAATCGAGAGACAGCTG 382
Oy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerIysGluLeuGluAlaAlaGlnAla 120
Db 383 ACCCGGTGGCGAGAGAGAGCGGCGGCGGCGGCTGTCCAAAGAGCTGCAGCGGCGAGGCC 442
Oy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrrArgGlyGluVal 140
Db 443 CGGTGGGTGCGGACATGAGAGACGTCGCGACGCGCTGTGCACATCCAGCAGCAGGTG 502
Oy 141 GluAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 503 CAGGCGCATCTGGCCAGAGTACCGGAGCTGGCGGCGGCGCTCCCTCCACCTGGCGC 562
Oy 161 LysLeuCysIysArgLeuLeuArgAspAlaAspAspLeuGluLysArgLeuAlaValTyr 180
Db 563 AAGCTTCGCAACGCGCTCTCCGCGATGTGATGACCTCAGAGAGCGCTTGCAGAGTGT 622
Oy 181 GlnAlaGlyAlaArgGluGluIysAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 623 CAGGCGGGGCGCGGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 682
Oy 201 ProLeuValGluGluGluArgValArgAlaAlaThrValGlySerLeuAlaGluGlnPro 220
Db 683 CCCCTGGGTGACAGAGGCGCGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 742
Oy 221 LeuGluGluArgAlaGluAlaTrrPgluIysArgLeuArgAlaArgMetGluGluMetGly 240
Db 743 CTTCAGAGCGGCGGCGGCGGCTTGGGTGAGCGGCTTCCGCGCAGCATGAGAGATGGGC 802
Oy 241 SerArgThrArgAspArgLeuAspGluValIysGluGluValAlaGluValAlaArgAlaIys 260
Db 803 ACCCGGAGCGCGGAGCGCGCTGTGACAGGTGAAGAGACAGTGGGAGGTGGCGCGCAAG 862
Oy 261 LeuGluGluGluAlaGluGluIleArgLeuGluAlaGluAlaPheGluAlaArgLeuIys 280
Db 863 CTGGAGGAACAGCGCCAGAGATAGCTGACGCGCGGAGCGCTTCCAGGCGCGCTCAAG 922
Oy 281 SerTrrPhegluProleuValgluAspMetgluArggluTrrPalagluLeuValgluIlys 300
Db 923 AGCTGTTGACCCCTGGTGGAGACATATGCAGCGCCAGTGGGCTGGTGGAGAAG 982
Oy 301 ValgluAlaValgluIlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 983 GTGCAGGCTGCGGTGGCGCACAGCGCCCTGTGCCATGCACATATCAC 1033
RESULT 15
AF261279 AF261279 5491 bp DNA linear PRI 27-OCT-2000
LOCUS Homo sapiens apolipoprotein-E gene, complete cds.
DEFINITION AF261279
ACCESSION AF261279.1 GI:11034800
VERSION
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 5491)
AUTHORS Nickerson,D.A., Taylor,S.L., Fullerton,S.M., Weiss,K.M.,
Clark,A.G., Stengard,J.H., Salama,V., Boerwinkle,E. and Sing,C.F.
TITLE Sequence diversity and large-scale typing of SNPs in the human
apolipoprotein E gene
JOURNAL Genome Res. 10 (10), 1532-1545 (2000)
MEDLINE 20499366
PubMed 11042151
REFERENCE 2 (bases 1 to 5491)
AUTHORS Nickerson,D.A.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2000) Department of Molecular Biotechnology,
University of Washington, Box 357730, Seattle, WA 98195, USA
FEATURES
source 1..5491
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.2"
/contig="108"
/note="putative"
/rpt_family="MIR"
/rpt_type="dispersed"
73
/frequency="0.01"
/replace="t"
repeat_region
/note="putative"
/rpt_family="MIR"
/rpt_type="dispersed"
308
/frequency="0.01"
/replace="t"
satellite
/note="putative"
/rpt_type="tandem"
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
471
/frequency="0.01"
/replace="g"
variation
/frequency="0.01"
/replace="t"
variation
/frequency="0.01"
/replace="t"
variation
/frequency="0.22"
/replace="t"
variation
/frequency="0.07"
/replace="c"
/note="putative"
/rpt_type="tandem"
832
/frequency="0.45"
/replace="t"
variation
/product="apolipoprotein-E"
1163
/frequency="0.35"
/replace="c"
variation
/frequency="0.01"
/replace="a"
1575
/frequency="0.01"
/replace="t"
/note="APOE"
/codon_start=1
CDS
join(1878..1920,3013..3205,3786..4503)
/note="APOE"
/codon_start=1

/product="apolipoprotein-E"
/protein_id="AAG27089.1"
/db_xref="GI:11034801"
/translation="KKVLAALALVFLAGCCQAKVBOAVETEPEPELROOTEMOSGQRM
EIALGRFMDYLRLWNOTLSROVOEELISQVDELRLADETKKELKAKSELEQLT
VAEETRLRSKELQAAQRLQADNEDVCRLVQYGEVQAMIGOSTEELRVLASHLR
KLKRLRLRDADLQRLAVYQAGARGARGLSAIRELGPLVEGRRKRAATVGSILAG
QPLQERAQAMGGERLARMEGMSRTDRIDEVKEQVAVERARLLEQDAQIRLQAEAFQ
ARLKSWEFEPLVEDMOROMAGLVEKVOAAVGTSAAPVPSDNH"
1998
/frequency="0.10"
/replace="a"
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
2440
/frequency="0.21"
/replace="a"
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
2907
/frequency="0.01"
/replace="g"
variation
/frequency="0.01"
/replace="c"
repeat_region
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
3673
/frequency="0.01"
/replace="g"
variation
/frequency="0.01"
/replace="g"
variation
/frequency="0.01"
/replace="cl"
variation
/frequency="0.14"
/replace="c"
variation
/frequency="0.01"
/replace="t"
variation
/frequency="0.07"
/replace="t"
repeat_region
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
5065..55476
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
5229~5230
/frequency="0.03"
/replace="gg"
variation
/frequency="0.07"
/replace="t"
variation
/frequency="0.40"
/replace="g"
variation
/frequency="0.13"
/replace="a"
5361
/frequency="0.06"
/replace="c"
BASE COUNT 1029 a 1663 c 1603 g 1196 t
ORIGIN

Alignment Scores:

Pred. No.:	6,8e-74	Length:	5491
Score:	1388.50	Matches:	303
Percent Similarity:	60.60%	Conservative:	0
Best Local Similarity:	60.60%	Mismatches:	3
Query Match:	87.38%	Indels:	14
DB:	9	Gaps:	1

US-09-827-854-18 (1-317) x AF261279 (1-5491)

```

QY 12 PheLeuAlaGlycysGlnAlaLysValGluGlnAlaValGluThrLupProGluProGlu 31
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3003 TTCACACAGAGATGCCAGGCGCAAGCGTGAGCAAGCGTGAGAGACAGAGCGGAGCCGAG
    3062

QY 32 LeuArgGlnGlnThrGluTrpGlnSerGlyGlnArgTrpGlnLeuAlaLeuGlyArgPhe 51
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3063 CTGCGCGAGAGAGACCGAGGTGGAGCGCCAGCGCTGGAGACTGGAGCTGGCTGCTTT
    3122

QY 52 TrpAspTyrLeuArgTrpValGlnThrLeuSerGlyGlnValGlnGluLeuLeuSer 71
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3123 TGGGATTACTGGCTGGCTGGGTGGAGACACTGTCTGACAGGTGCGAGAGAGAGCTGCTCAGC
    3182

QY 72 SerGlnValThrGlnLeu----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3183 TCCAGGTACCCAGAGAACTGAGTGTGATGCTGCCATCTGCGCTTGACCTCTGTGTG
    3242

QY 78 ----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3243 GCGGCTATACCTCCCGAGGTCAGGTTCATTCTGCCCCCTGTGCTAAGTCTTGGGGG
    3302

QY 78 ----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3303 CCGGGTCTCTGTGTGTCTAGCTTCTCTCCATTCTGACTCTGCTGCTTACTGCTTC
    3362

QY 78 ----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3363 TGGAAATCTCTCTCAGCTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
    3422

QY 78 ----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3423 CGTCTGGCTCTGTCTGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
    3482

QY 78 ----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3483 CTCACGTGTGTGCCAGGCTGTCTGAACTCTGGGCTCAACGATCCTCCGCTCTCG
    3542

QY 78 ----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3543 CCTCCAAAGTGTGGGATTAGAGCATAGACCACTTGCCGCGCTCTAGCTCTCTCT
    3602

QY 78 ----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3603 TCGTCTGCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
    3662

QY 78 ----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3663 GCCTGCCCCGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
    3722

QY 78 ----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3723 CCCCATCCAGCCCTCTCTCCCGCTCCCACTGTGAGACACCTCCCGCTCTCGCGG
    3782

QY 79 -ArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlu 98
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3783 CAGGGCGCTGATGAGACGATGAAAGGCTTGAAGGCTTACAAATCGGAACCTGGAGGA
    3842

QY 98 uGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAla 118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3843 ACAACTGACCCCGCTGGCGAGAGACGCGGCGACGCTGTCCAGAGAGCTGACGCGCG
    3902

QY 118 aglnAlaArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGln 138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 3903 GCAGCCCGGCTGGCGCGACATGGAGAGACTGTGCGGCCCGCCCTGCTGACATACCGCGG 3962

QY 138 yGluValGlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerH 158
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3963 CAGAGTGCAGGCCATCTCGGCCAGAGACACAGAGAGCTCGGGTGCCTCGCTCCCA 4022

QY 158 sLeuArgLysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAl 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4023 CTTGCCAGAGCTCGTAAGAGGGCTCTCCGCCATTCGATGACCTGCAGAAAGCGCTGGC 4082

QY 178 aValTyrGlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaLeuArgLuar 198
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4083 AGGTACCAAGCGCGGGCGCGAGGCGCGCGAGCGCGCTTCAGCCATCCGAGGAGG 4142

QY 198 gLeuGlyProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaG 218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4143 CTTGGGGCCCCCTGTGTGAACAGGCGCGCTGTGCGGCGCCACTGTGGGCTCCTGGCCG 4202

QY 218 yGlnProLeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluG 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4203 CCAGCGGCTACAGAGACGCGGCCAGGCTGTGGGCGAGCGGCTCGCGCGGATGGAGGA 4262

QY 238 uMetGlySerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAr 258
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4263 GATGGGACAGCCGAGCCCGGAGCGCTGTGAGAGAGTGAAGAGACAGTGGCGAGGTGGG 4322

QY 258 gAlaLysLeuGluGluGlnAlaGlnGlnLeuArgLeuGlnAlaGlnAlaPheGlnAlaAr 278
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4323 CGCAGAGCTGGAGAGAGAGCCAGCAGATACCTGTGAGGCGGAGGCGCTTCCAGGCGCG 4382

QY 278 gLeuLysSerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuVa 298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4383 CCTCAAGAGCTGTGTGAGAGCCCTGTGTGAAGACATGACAGCGCACTGGGCGGCTGGT 4442

QY 298 lGluLysValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4443 GGAGAGGTGCAAGCTGTGGCGACAGCGCGCCCTGTGCTCCAGCGACATATCAC 4500

```

Search completed: March 14, 2003, 17:32:43
 Job time : 2314.23 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:08:17 : Search time 178.728 seconds

(without alignments)
3994.237 Million cell updates/sec

Title: US-09-827-854-18

Perfect score: 1589

Sequence: 1 MKVLMAALTVFLAGCOAKV.....VEKVOAAGTSAPVSDNH 317

Scoring table:

BLOSUM62	
Xgapop 10.0, Xgapext 0.5	
Ygapop 10.0, Ygapext 0.5	
Fgapop 6.0, Fgapext 7.0	
Delop 6.0, Delext 7.0	

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlp
-O=/cgn2.1/USPto_spool/US09827854/runat_11032003_101609_27476/app_query.fasta_1.3576
-DB=N_Geneseq_101002 -QFMT=fastap -SUFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptco -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854 -CGCN 1.1.1201 -runat_11032003_101609_27476 -NCPU=6 -ICPU=3
-NO_XIPXY -NO_MMAP -LARGEQUERY -NEG1SCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :

1:	N_Geneseq_101002.*
2:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24:	/SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1589	100.0	1156	24	AAD22051	Human apolipoprotein E
2	1577	99.2	954	24	AAD26035	Human apolipoprotein E
3	1577	99.2	1110	7	AA060409	Human apolipoprotein E
4	1577	99.2	1147	22	AB83113	Human apolipoprotein E
5	1577	99.2	1156	22	AA084315	Human apolipoprotein E
6	1577	99.2	1156	24	AAD22048	Human apolipoprotein E
7	1573	99.0	1156	24	AAD22052	Human apolipoprotein E
8	1569	98.7	1156	22	AA084314	Human apolipoprotein E
9	1569	98.7	1156	24	AAD22049	Human apolipoprotein E
10	1568	98.7	1157	17	AA060957	Human apolipoprotein E
11	1568	98.7	1157	24	AB095746	Human apolipoprotein E
12	1568	98.7	1157	24	AB064514	Human apolipoprotein E
13	1568	98.7	1157	24	AB064510	Human apolipoprotein E
14	1567	98.6	1110	6	AA050450	Human apolipoprotein E
15	1565	98.5	1156	22	AA084316	Human apolipoprotein E
16	1565	98.5	1156	24	AAD22047	Human apolipoprotein E
17	1562	98.3	1156	24	AAD22050	Human apolipoprotein E
18	1473.5	92.7	1279	22	AA022437	Human apolipoprotein E
19	1454.5	91.5	1107	19	AA075756	Human apolipoprotein E
20	1388.5	87.4	9360	24	AB031915	Human apolipoprotein E
21	1388.5	87.4	10716	24	AAD26034	Human apolipoprotein E
22	1383	87.0	3805	20	AA020524	Human apolipoprotein E
23	1383	87.0	3805	20	AA020526	Human apolipoprotein E
24	1361.5	85.7	10716	24	AAD26108	Human apolipoprotein E
25	1162.5	73.2	965	24	AAD32081	Human apolipoprotein E
26	1162.5	73.2	5617	24	AAD32077	Human apolipoprotein E
27	1162.5	73.2	6026	24	AAD32075	Human apolipoprotein E
28	1130	71.1	1126	19	AA028159	Human apolipoprotein E
29	977	61.5	936	15	AA069101	Human apolipoprotein E
30	977	61.5	936	17	AA018070	Human apolipoprotein E
31	976	61.4	660	18	AA069792	Human apolipoprotein E
32	939	59.1	597	17	AA018068	Human apolipoprotein E
33	930	58.5	597	15	AA069099	Human apolipoprotein E
34	879	55.3	1381	22	AA022673	Human apolipoprotein E
35	877.5	55.2	786	21	AA018114	Human apolipoprotein E
36	670	45.2	600	20	AA089595	Human apolipoprotein E
37	651	41.0	407	24	AB034238	Human apolipoprotein E
38	625	39.3	478	24	AB095994	Human apolipoprotein E
39	625	39.3	478	24	AB062679	Human apolipoprotein E
40	625	39.3	478	24	AB067340	Human apolipoprotein E
41	594	37.4	499	22	AA030349	Human apolipoprotein E
42	506	31.8	330	12	AA011980	Human apolipoprotein E
43	504	31.7	405	21	AA021339	Human apolipoprotein E
44	485	30.5	345	22	AA098479	Human apolipoprotein E
45	411	25.9	260	21	AAA040342	Human apolipoprotein E

ALIGNMENTS

RESULT 1
AAD22051
ID AAD22051 standard: DNA: 1156 BP.

AC AAD22051: 12-FEB-2002 (first entry)

DE Human apolipoprotein E (apoE) allele, apoE2* DNA.

KW Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;

KW hypertriglyceridaemia; low density lipoprotein; LDL; ds.

OS Homo sapiens.

XX key Location/Qualifiers
FH 61..1014
FT CDS /*tag= a

FT sig.peptide /product= "Human apoE allele, apoE2*" 61..114
FT /*tag= b 115..1011
FT mat.peptide /*tag= c
FT /product= "Mature human apoE allele, apoE2*" 18-OCT-2001.
XX WO200177136-A1.
XX 18-OCT-2001.
PD 06-APR-2001. 2001MO-US11358.
XX
PF 06-APR-2000. 2000US-0544386.
XX 06-APR-2000. 2000US-0544386.
PR 04-OCT-2000. 2000US-0679088.
PR 05-APR-2001. 2001US-0827854.
XX
PA (KOSP-) KOS PHARM. INC.
PA (UYBO-) UNIV BOSTON.
XX
PI Zannits VI, Kypreos KE;
XX
PI
XX
DR WPI: 2002-010885/01.
P-PDB: AAE13297.
XX
XX
PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT cholesterol, delaying the onset of or treating atherosclerosis in
PT mammal, without inducing hypertriglyceridemia -
XX
XX
PS Claim 14: Page 82: 91pp: English.
XX
CC The present sequence is a human apolipoprotein E (apoE)
CC allele, apoE2* DNA. The apoE lipoproteins are useful for
CC lowering cholesterol, delaying the onset of atherosclerosis,
CC treating or reversing atherosclerosis without inducing
CC hypertriglyceridaemia, in a mammal lacking an endogenous,
CC normally functioning apoE gene or low density lipoprotein (LDL)
CC receptor or is at risk for developing atherosclerosis due to
CC accumulation of lipoprotein remnants in the bloodstream or having
CC a defect in remnant removal.
XX
SQ Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other:

Alignment Scores:
Pred. No.: 3.49e-116 Length: 1156
Score: 1589.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

US-09-827-854-18 (1-317) x AAD22051 (1-1156)
QY 1 MetLysValLeuTrpAlaIleuValThrPheLeuAlaGlyCysGlnAlaIleVal 20
DB 61 ATAAAGGTTCTGGGGCTGGGTGGTGCATTCCTGGCAGATGCAAGGCGT 120
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 121 GACCAAGCGGTGGAGACAGACGCCGAGCTCGCCAGACAGCCAGTGGCAGAGC 180
QY 41 GlyGlnAlaArgTrpGluLeuAlaIleuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 181 GGCACACCGCTGGGAACTGGCACTGGGCTCTTTGGGATTACTGCGTGGTGACAGACA 240
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 241 CTGTCTGACCAAGTGCAGAGAGAGCTGCTCAGCTCCACAGTCCACAGAACTGAGGGCG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnGlnLeu 100
DB 301 CTCATGACGAGACCATGTAAGAGATTGAAGGCTTACAATCGGAATCGAGAGAACACTG 360

QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIleAla 120
DB 361 ACCCGGTGGCGAGAGAGACGCGGCACGCTCTCCAGAGAGCTGCAGGCGCGCAGGCC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnThrArgGlyGluVal 140
DB 421 GGGCTGGGCGGACATGGAGACGTGTGGGCGGCTGGTGCAGTACCGCGGAGGCTG 480
QY 141 GluAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCGATGCTGGCGACAGACACGAGAGCTCGGGGTGGCTCCCTCCACCTGCGCG 540
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 541 AACCTGTATACCGCTCCCTCCCGATGCCGATGACCTGAGAACGCGCTGGCAGCTGAC 600
QY 181 GlnAlaGlyAlaArgGluGluGluArgGlyLysSerAlaIleArgGluArgGluGly 200
DB 601 CAGCGCGGGCGCGGAGGGCGCGAGCGGCGCTCAGCGCCATCCGCGAGCGCTGGGG 660
QY 201 ProLeuValGluGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 661 CCCCTGGTGGAAACAGGGCGCGCTGGCGGCGCGCACCTGGGCTCCCTGGCGCGCAGCG 720
QY 221 LeuGlnGluArgAlaGlnAlaThrGlyGluArgLeuArgAlaArgMetGluMetGly 240
DB 721 CTACAGAGAGCGGCGCCAGCGCTGGCGGCGAGCGGCTCGCGCGCGATGAGAGATGGCG 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 781 AGCGGAGCCCGACCGCTCGGACGAGTGAAGACAGCAGGTGGAGGTGGCGGCAAG 840
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 841 CTGAGAGAGAGAGCCAGCAGCAGATAGCCTGACGCGGAGCGCTCCAGGCGCGCTCAAG 900
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 901 AGCTGTTTGAGCGCCCTGCTGGTGAAGACATGACGCGCAGTGGGCGCGCTGGTGGAGAG 960
QY 301 ValGlnAlaIleValGluThrSerAlaIleProValProSerAspAsnHis 317
DB 961 GTGCAAGCTGCGCTGGGCAACGCGCGCCCTGTGCCAGGACATAC 1011

RESULT 2
AAD26035
ID AAD26035 standard; cDNA; 954 BP.
XX
XX AAD26035;
AC
XX
XX 26-MAR-2002 (first entry)
DT
DT
DE Human apolipoprotein E (APOE) cDNA.
XX
XX
KW Human; anti-lipemic; neuroprotective; noctropic; genetic variant; APOE;
KW apolipoprotein E; haployping; familial dysbetalipoproteinemia; therapy;
KW genotyping; type III hyperlipoproteinemia; Alzheimer's disease; SNP;
KW atherosclerosis; single nucleotide polymorphism; chromosome 19q13.2; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..954
FT /*tag= a
FT /product= "Human APOE protein"
FT /replace (13, C)
FT /*tag= a
FT /standard.name= "Single nucleotide polymorphism (SNP)"
FT /replace (31, G)
FT /*tag= b
FT /standard.name= "Single nucleotide polymorphism (SNP)"
FT /replace (364, A)
FT /*tag= c


```

FT      variation      /standard_name="Single nucleotide polymorphism (SNP)"
FT      replace (388; C) /tag= d
FT      variation      /standard_name="Single nucleotide polymorphism (SNP)"
FT      replace (498; G) /tag= e
FT      variation      /standard_name="Single nucleotide polymorphism (SNP)"
FT      replace (526; T) /tag= f
FT      variation      /standard_name="Single nucleotide polymorphism (SNP)"
FT      replace (622; A) /tag= g
FT      variation      /standard_name="Single nucleotide polymorphism (SNP)"
PN      WO200179234-A2.
XX      25-OCT-2001.
XX      16-APR-2001; 2001WO-US12303.
XX      14-APR-2000; 2000US-197188P.
XX      (GENA-) GENAISSANCE PHARM INC.
XX      Chou JY, Klem SE, Koshy B, Lee HH;
XX      MPI: 2002-075064/10.
XX      P-PSDB: AAE15158.
XX      Genotyping human apolipoprotein gene of individual for determining
XX      haplotype of individual, involves determining identity of nucleotide
XX      pair at specific polymorphic sites for two copies of gene -
XX      Claim 26; Fig 2; 78pp: English.
XX      The patent discloses novel genetic variants of human apolipoprotein
XX      E (APOE) gene. The invention also relates to compositions and methods
XX      for haplotyping and/or genotyping the APOE gene. The haplotyping
XX      methods of the invention are useful for improving the efficacy and
XX      reliability of several steps in the discovery and development of
XX      drugs for treating diseases associated with APOE activity, e.g.
XX      familial dysbetalipoproteinemia, type III hyperlipoproteinemia,
XX      atherosclerosis, and Alzheimer's disease. They are useful to validate
XX      APOE as a candidate agent for treating a specific condition or disease
XX      predicted to be associated with APOE activity and in the design of
XX      clinical trials of candidate drugs for treating a specific condition
XX      or disease predicted to be associated with APOE activity. Genotyping
XX      or haplotyping methods are useful to screen for compounds targeting
XX      APOE to treat a specific condition or disease associated with APOE
XX      activity. The present sequence is a cDNA encoding human APOE protein.
XX      APOE gene is located on chromosome 19q13.2.
SQ      Sequence 954 BP; 168 A; 290 C; 377 G; 119 T; 0 other;
Alignment Scores:
Pred. No.: 2,47e-115
Score: 1577.00
Percent Similarity: 99.68%
Best Local Similarity: 99.68%
Query Match: 99.24%
DB: 24 Gaps: 0
US-09-827-854-18 (1-317) x AAD26035 (1-954)
QY      1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB      1 ATGAAGGTTCTGTGGGCTCGTTCGTCACATTCCTGGAGATGCCAGGCCAAGGTG 60
QY      21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB      61 GAGCAAGCGGTGGAGACAGCCGCGAGCCGAGCTCGCCAGCAGCCAGTGGCAGAGC 120
QY      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60

```

```

DB      121 GGCACAGCGCTGGACATGCGACTGGGTCCTTTGGAGTTACTGCGGTGGCAGACA 180
QY      61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB      181 CTGTGTGACGACGATGAGAGAGAGGCTGCTCACCCTCCAGTACCAGAACTGAGAGCGG 240
QY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
DB      241 CTGATGGACGACACCATGATGAGAGCTTGAAGGCTTACAAATCGGAAGTGGAGACACTG 300
QY      101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB      301 ACCCGCGTGGCGAGAGAGACGCGGCGACGCTGTCCAAAGACTGACAGCGCGCGAGGCC 360
QY      121 ArgLeuGlyAlaAspMetLysValLysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB      361 CGGCTGGGCGGACATGTAGAGACGTGCGGCGGCTGTGTGATGACCGCGGAGGTG 420
QY      141 GlnAlaMetLeuGlyGlnSerThrGluGluArgValArgLeuAlaSerHisLeuArg 160
DB      421 CAGGCCATCTCGCCAGACACCGAGAGCTGCGGCTCGCTCCACTCCACTGCGC 480
QY      161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaLysTyr 180
DB      481 AAGCTCGTAAAGCGGCTCTCCGCGATGCGGATGACTGACAGAGCGCTGGCAGGTAC 540
QY      181 GlnAlaGlyAlaArgGluGlnAlaGluArgGlyLeuSerAlaThrArgGluArgLeuGly 200
DB      541 CAGCGCGGGGCGCGCGAGGCGCGCGCGCTCCAGCGCATCCGCGCAGCGCGCTGGGG 600
QY      201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB      601 CCCCTGTGGAACAGGCGCGCTGCGGCGCCACTGTGGCTCCTGGCCGCGCAGCCG 660
QY      221 LeuGlnGluArgAlaGlnAlaThrGlyGluArgLeuArgAlaArgMetGluMetGly 240
DB      661 CTACAGAGACGCGGCGCGAGCTGCGGCGAGCGGCTCGCGCGGATGAGAGATGGGC 720
QY      241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB      721 ACCCGGACCCGCGACCGCTGTGACGAGGTGAAGAGCAGGTGGCGAGGTGCGCGCAAG 780
QY      261 LeuGlnGluGlnAlaGlnGlnThrArgLeuGlnGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB      781 CTGGAGGACAGGCGCCAGAGATACCTGCGAGGCCGAGGCTTCCAGGCCCGCTCAAG 840
QY      281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB      841 ACCTGTTGAGAGCCCTGTGTGGAAGACATGACAGCCCACTGGGCGCGGTGGTGGAGAG 900
QY      301 ValGlnAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB      901 GTTCAGGCTGTGGTGGCACACAGCGCGCCCTGTGTCACGAGCAATATCAG 951
RESULT 3
AAN60409
ID      AAN60409 standard; DNA; 1110 BP.
AC      AAN60409;
XX      01-JAN-1980 (first entry)
XX      Human apolipoprotein-E.
XX      Apolipoprotein-E; hyperlipidemia; atherosclerosis; ss.
XX      Homo sapiens.
XX      Key      Location/Qualifiers
XX      FT      CDS      15..968
XX      FT      /*tag= a
XX

```


PN A08547513-A.
 XX
 PD 24-APR-1986.
 XX
 PF 17-SEP-1985; 85AU-0047513.
 XX
 PK 11-JUN-1985; 85UP-0126989.
 XX
 PA (MITU) MITSUBISHI CHEM IND KK.
 XX
 DR WPI: 1986-150217/24.
 P-PSDB; AAP60507.
 XX
 PT New DNA sequence coding for human apolipoprotein-E - and
 expression vectors and transformed cells contg. it
 XX
 PS Disclosure; Fig 2; 45pp; English.
 CC The encoded protein is used to treat subjects who are deficient in
 CC apolipoprotein-E (or who produce abnormal forms of this molecule)
 CC and therefore are likely to suffer from hyperlipidemia, resulting in
 CC arteriosclerosis. It can also be used to raise antisera for
 CC detecting the protein deficiency or production of abnormal forms.
 XX

SO Sequence 1110 BP; 198 A; 354 C; 415 G; 143 T; 0 other;

Alignment Scores:

Pred. No.:	2,93e-115	Length:	1110
Score:	1577.00	Matches:	316
Percent Similarity:	99.68%	Conservative:	0
Best Local Similarity:	99.68%	Mismatches:	1
Query Match:	99.24%	Indels:	0
DB:	7	Gaps:	0

US-09-827-854-18 (1-317) x AAN60409 (1-1110)

QY 1 MetLysValLeuThrPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIleVal 20
 Db 15 ATGAAGGTTCTGTGGTGGCTGCTGCTGCATCTCTGGCAGAGATGCCAGGCCAAGTG 74
 QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnThrGlnSer 40
 Db 75 GAGCAACGGGTGGAGAGAGACCCGAGCTGGCCGACGACGACGAGTGGCAAGC 134
 QY 41 GlnGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 Db 135 GGCAGGCGTGGGAATGCACTGGCTGCTTTGGATTACTGCGCTGGGTGCAGACA 194
 QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 Db 195 CTGCTGAGCAGGTGGCAGGAGAGCTCTCAGCTCCAGGTCACCGAAGTGAAGGGCG 254
 QY 81 LeuMetAspGlnThrMetLysGlnLeuLeuAlaTyrLysSerGlnLeuGlnGlnLeu 100
 Db 255 CTGATGACGACGACCATGAAGAGCTTGAAGCCCTACAAATTCGACACTGGAAGAACTG 314
 QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
 Db 315 ACCCCGCTGGCGGAGGAGACGGGGCAGCGCTGTCCAAAGAGTGCAGCGCGGCAGAGCC 374
 QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnThrArgGlnVal 140
 Db 375 CGGCTGGGGCCCGACATGAGAGAGCTGTGGCGCCCTGTGTGACGTACCCCGCGAAGTG 434
 QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 435 CAGGCCATGCTCGGCCAGACACCGAGAGCTGGGGTGGCCCTCGCTCCACACTGCGC 494
 QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 Db 495 AAGCTGCGTAAAGCGGCTCTCCGAGATGCGATGACTCTGCAAGAGCGCTGGCAGTGTAC 554
 QY 181 GlnAlaGlnAlaArgGlnGlnGlnAlaGlnArgGlyLeuSerAlaIleArgGlnValGlnGly 200

Db 555 CAGCCGGGGCCCGGAGGGCGCCGAGCGCGCTCAGCGCATCCCGAGCGCTGGG 614
 QY 201 ProLeuValGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 615 CCCCTGTTGGAACAGGGCCCGCTGCGGGCCGCCACTGTGGCTCTCTGGCCGCGACCG 674
 QY 221 LeuGlnGlnAlaGlnAlaTyrGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
 Db 675 CTACAGAGAGGGGCCAGCGCTGGGCGAGCGCTGCGCGCGCGATGAGAGATGGCG 734
 QY 241 SerArgThrArgAspAlaGlyLeuAspGlyValLysGlnGlnValAlaGlnValArgAlaLys 260
 Db 735 AGCCGGACCCCGCGCCCTGGAGAGGTGAAGAGAGGTGGCGGAGGTGGCGCAAG 794
 QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 Db 795 CTGGAGAGACGAGCCCGCAGCATTCAGCTTCAGAGCCGAGGCTTCAGAGCCCGCTCAAG 854
 QY 281 SerTrpPheGlnProLeuValGlnAspMetGlnArgGlnThrPAlaGlyLeuValGlnLys 300
 Db 855 AGCTGTTTCAGGCCCTGTGTGAAGATGATGACAGCGCCAGTGGCGCGGCTGTGTGAGAG 914
 QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 915 GTGCAGGCTGCGCTGGGACACGCGCCCGCTGTGCCAGCGACATTCAC 965

RESULT 4

ABAB3113 standard; DNA; 1147 BP.

ABAB3113;

08-FEB-2002 (first entry)

Apollipoprotein E ovarian tumour marker gene, SEQ ID NO:63.

KW Ovarian tumour marker gene; human; overexpression; upregulation;
 KW epithelial tumor; cancer; diagnosis; prognosis; disease monitoring;
 KW identification; serous cystadenoma; borderline serous tumor;
 KW serous cystadenocarcinoma; mucinous cystadenocarcinoma;
 KW mucinous cystadenoma; borderline mucinous tumor; endometrioid carcinoma;
 KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;
 KW adenofibroma; Brenner tumor; serial analysis of gene expression; SAGE;
 KW immune response pathway; cell proliferation regulation; protein folding;
 KW membrane localised; secreted; therapeutic target; cytostatic;
 KW gene therapy; vaccine; ds.

OS Homo sapiens.

PN WO200175177-A2.

PD 11-OCT-2001.

PF 03-APR-2001; 2001WO-US10947.

PR 03-APR-2000; 2000US-194336P.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;

DR WPI: 2001-626450/72.

DR P-PSDB; ABB50287.

PT detecting and identifying ovarian tumor, identifying increased risk for

PT developing ovarian cancer, and determining effectiveness of ovarian

PT cancer treatment, by measuring expression level of ovarian tumor marker

PS gene -
 Claim 23; Page 105-106; 140pp; English.

XX The invention relates to methods for diagnosing and prognosing ovarian

CC tumours in an individual via the detection and measurement of the
 CC expression of ovarian tumour marker genes (ABA83081-ABA83122, ABA83180,
 CC ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179,
 CC ABA83181 and ABA83183). The methods of the invention are useful for
 CC detecting an ovarian tumour in a patient, for identifying an individual
 CC at increased risk for developing ovarian cancer, in prognostic tests for
 CC assessing the relative severity of ovarian cancer, in tests for
 CC monitoring a patient in remission from ovarian cancer and in tests for
 CC monitoring disease status in a patient being treated for ovarian cancer.
 CC The methods can additionally be used to identify a particular tumour as
 CC being an ovarian tumour (i.e., an epithelial ovarian tumour selected from
 CC serous cystadenoma, borderline serous tumour, serous cystadenocarcinoma,
 CC mucinous cystadenoma, borderline mucinous tumour, mucinous
 CC cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma,
 CC clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner
 CC tumour. The ovarian tumour marker genes of the invention were identified
 CC using SAGE (Serial analysis of gene expression) and were found to be
 CC overexpressed in a broad variety of ovarian epithelial tumour cells
 CC relative to normal ovarian epithelial cells. The marker genes are
 CC implicated in immune response pathways, in the regulation of cell
 CC proliferation and in protein folding, and many of these are membrane-
 CC localised or secreted. In addition to their use as diagnostic and
 CC prognostic markers, the ovarian tumour marker genes or their encoded
 CC proteins may be used as therapeutic targets for the treatment and
 CC prevention of ovarian cancer. Sequences ABA83081-ABA83122, ABA83180,
 CC ABA83182 and ABA83184 represent the ovarian tumour marker genes of
 CC the invention.

XX Sequence 1147 BP; 210 A; 365 C; 425 G; 147 T; 0 other;

Alignment Scores:

Pred. No.: 3,04e-115
 Score: 1577.00
 Percent Similarity: 99.68%
 Best Local Similarity: 99.68%
 Query Match: 99.24%
 DB: 22

Length: 1147
 Matches: 316
 Conservative: 0
 Mismatches: 1
 Indels: 0
 Gaps: 0

US-09-827-854-18 (1-317) x ABA83113 (1-1147)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 46 ATGAAGGTTCTGTGGGCTCGTGTGTCGTCACATCTCTGGACAGATGCCAGGCAAGTG 105
 QY 21 GlnGlnAlaValGlnThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 106 GACCAAGCGGTGGAGACAGACCGGAGCCGAGACTCGCCAGACAGACAGTGGCAGAGC 165
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 DB 166 GGCACAGCGTGGAACTGGCAGTGGCTTTTGGATTACCTGGCTGGGTCAGACA 225
 QY 61 LeuSerGluGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 DB 226 CTGTGTGACAGAGTCAGAGAGCTGCACGCTCCAGATCCACCAAGAACTGAGGGG 285
 QY 81 LeuMetAspGlnThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu 100
 DB 286 CTGATGGAGAGACCATGTAAGAGTTGAAGGCTTCAAAATCGGAATGGAGGAACAAC 345
 QY 101 ThrProValAlaGlnGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 346 ACCCGGTTGGCGAGAGACCGCGGCTCTCCAAAGAGCTGCAAGCGCGCGAGGCC 405
 QY 121 ArgLeuGlyAlaAspMetGluAspValLysGlyArgLeuValGlnTrpArgGlyLysVal 140
 DB 406 CGGCTGGCGCGACATGAGAGACGTCGCGCGCCCTGTGCAGTAACCGCGCGAGTG 465
 QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 466 CAGGCGCATCTCGGCGAGACACCGGAGAGCTGGGGTGGCTCCCTCCACCTCGCGC 525
 QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180

DB 526 AAGCTCGTAAGAGGCTCTCCGCGATGCGATGACCTGCAGAAAGCGCTGGCAGTGAC 585
 QY 181 GlnAlaGlyAlaArgGluGluValGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 586 CAGGCGGGGCGCGCAGAGGCGCGGCGGCTTACGCGCATCCGACAGCGCCCTGGGG 645
 QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
 DB 646 CCCCTGGTGAACAGAGGCGCGGCGCGCCCTCTGTGGCTCTCTGGCGGCGCAGCG 705
 QY 221 LeuGlnGluArgAlaGlnAlaIleTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
 DB 706 CTACAGAGACCGGCGCAGGCTGTGGCGAGCGGCTGCGCGCGGATGGAGAGATGGGC 765
 QY 241 SerArgTrpArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
 DB 766 AGCCGAGACCGCGACCGCTGTGACGAGGTGAAGAGACAGTGGCGAGGTGGCGCAAG 825
 QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 DB 826 CTGGAGAGACAGCGCCAGAGATACGCTGTGAGGCGGAGGCTTCCAGGCGCGCTCAAG 885
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 DB 886 AGCTGTGTGAGAGCCCTGTGTGGAAGACATGACAGCGCAGTGGCGCGGCTGTGAGAA 945

RESULT 5

AAF84315
 ID AAF84315 standard; cDNA; 1156 BP.

AAF84315;

21-JUN-2001 (first entry)

Human ApoE3 coding sequence.

Human; ApoE3; Alzheimer's disease; arteriosclerosis; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 61..1014

FT /tag= a

FT /product= "Human ApoE3"

JP2001017028-A.

23-JAN-2001.

28-APR-2000; 2000JP-0128919.

06-MAR-1999; 99JP-0125647.

(MITU) MITSUBISHI CHEM CORP.

WPI: 2001-285406/30.

P-PSDB: AAB80997.

New apoe humanized mammalian cell useful for screening for agents

arteriosclerosis -

Disclosure: Page 13-14; 22pp; Japanese.

The present invention relates to, an ApoE humanised mammalian cell. The present sequence is the coding sequence for human ApoE3, which was used in the method of the present invention. The ApoE humanised mammalian cell can be used for screening for agents useful for treating or preventing

CC Alzheimer's disease and arteriosclerosis.

XX Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other;

Alignment Scores:

Pred. No.:	3,07e-115	Length:	1156
Score:	1577.00	Matches:	316
Percent Similarity:	99.68%	Conservative:	0
Best Local Similarity:	99.68%	Mismatches:	1
Query Match:	99.24%	Indels:	0
DB:	22	Gaps:	0

US-09-827-854-18 (1-317) x AAF84315 (1-1156)

```

OY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 61 ATGAAGGCTTGTGGCTGGCTGGTGTGATTCCTGAGAGATTCAGGCCAAGAGT 120
OY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGlnThr 40
DB 121 GAGCAAGCGGTGAGACAGACCGGAGCCGAGCTGGCCAGACAGACCGAGTGCAGAGC 180
OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB 181 GGGCAGCGCTGGGACACTGGCTGGCTTTGGATTACCTGGCTGGTGCAGACA 240
OY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 241 CTGCTGAGCAGGTGCAGAGAGAGCTCTCAGCTCCAGAGTCAACCCAGGAAGTGAAGGCG 300
OY 81 LeuMetLyspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
DB 301 CTGATGGACAGACCAAGAGAGATTGAAGCCCTACAAATCGGAACGGAGGAACAACGTG 360
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 361 ACCCGCGTGGCGAGAGACCGGCGCACGCTGTCCAAAGACTGCGAGCGCGCAGAGCC 420
OY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 421 CGGCTGGCGCGGACATGAGAGAGCTGTGGCGCGCTGTGCAGTACCGCGGAGAGGTG 480
OY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCGATGCTCGGCGACAGACCGAGAGCTCGGGTGGCTCGCTCCACCTCCGCGC 540
OY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 541 AACCTGGTAAAGGGGCTCCCGCATGCCGATGACTCGAGAAGCCCTGGCAGATGAC 600
OY 181 GlnAlaGlyAlaArgGluGlyAlaGlyArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 601 CAGGCGGCGCGCGGAGCGCGCGAGCGCGGCTCAGCCATCCGCGAGCGCCCTGGGG 660
OY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 661 CCCCTGGTGAAGAGGCGCCCGCGCGCGCCACATGTGGGCTCCCTGGCGCGCCAGCCG 720
OY 221 LeuGlnGluArgAlaGlnAlaTyrGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
DB 721 CTACAGAGAGCGGCGCCAGGCTTGGGGAGCGGCTGGCGCGGATGGAGGAGATGGGC 780
OY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 781 AGCGGAGACCGCGACCGCGCGAGCGAGTAAAGAGACAGGTGGCGGAGTGGCGGCCAAG 840
OY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 841 CTGAGAGAGAGCGCCAGCAGATACCGCTGCAGCGCGAGCGCTTCCAGGCGCCGCTCAAG 900
OY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 901 AGCTGTGTCAGGCCCTGTGTGAAGATGCACGCGCCAGTGGCGGCGGTGTGTGAAGAG 960

```

```

OY 301 ValGlnAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 961 GTGCAGGCTGCGGTGGCGACAGCGCGCCCTGTGCCAGCAGCATTCAC 1011
RESULT 6
AAD22048
ID AAD22048 standard; DNA: 1156 BP.
XX
AC AAD22048;
XX
DT 12-FEB-2002 (first entry)
XX
DE Human apolipoprotein E (apoE) isoprotein, apoE3 DNA.
XX
KW Human; apolipoprotein E; apoE; cholesterol; arteriosclerosis;
KM hypertriglyceridaemia; low density lipoprotein; LDL; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 61..1014
FT /tag= a
FT /product= "Human apoE isoprotein, apoE3"
FT sig_peptide 61..114
FT /tag= b
FT mat_peptide 115..1011
FT /tag= c
FT /product= "Mature human apoE isoprotein, apoE3"
XX
PN WO200177136-A1.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001MO-US11358.
XX
PR 06-APR-2000; 2000US-0544386.
PR 04-OCT-2000; 2000US-0679088.
PR 05-APR-2001; 2001US-0827854.
XX
PA (KOSP-) KOS PHARM INC.
PA (UYBO-) UNIV BOSTON.
XX
PI Zannis VI, Kypros KE;
XX
DR MPI: 2002-010885/01.
DR P-PDB; AAE13294.
XX
PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT cholesterol, delaying the onset of or treating arteriosclerosis in
PT mammal, without inducing hypertriglyceridaemia.
XX
PS Claim 14; Page 81; 91pp; English.
XX
CC The present sequence is a human apolipoprotein E (apoE)
CC isoprotein, apoE3 DNA. The apoE lipoproteins are useful for
CC lowering cholesterol, delaying the onset of arteriosclerosis,
CC treating or regressing arteriosclerosis without inducing
CC hypertriglyceridaemia, in a mammal lacking an endogenous,
CC normally functioning apoE gene or low density lipoprotein (LDL)
CC receptor or is at risk for developing arteriosclerosis due to
CC accumulation of lipoprotein remnants in the bloodstream or having
CC a defect in remnant removal.
XX
SO Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other;

```

Alignment Scores:

Pred. No.:	3,07e-115	Length:	1156
Score:	1577.00	Matches:	316
Percent Similarity:	99.68%	Conservative:	0
Best Local Similarity:	99.68%	Mismatches:	1
Query Match:	99.24%	Indels:	0
DB:	24	Gaps:	0

US-09-827-854-18 (1-317) x AAD22048 (1-1156)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 61 ATGAAGGTTCTGTGGGCTGCTGTGTCACATTCTTGCGAGATGCGCAGGCAAGTG 120
 QY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 121 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTGCGCAGACAGACCGAGTGGCAGAGC 180
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 DB 181 GGCCAGCGCTGGAGTGGACACTGGCTGCTTTGGGATGATCTGGCTGGGTGGCAGACA 240
 QY 61 LeuSerGluGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB 241 CTGTCTGACAGTGCAGAGAGAGTGTCTCAGCTCCAGCTCACCAGGACCTGAGAGCGG 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu 100
 DB 301 CTGATGAGACGACACCATGAGAGGTTGAGAGGCTTACAAATCGGAATGGAGTGAAGACACTG 360
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 361 ACCCGGTGGCGAGAGAGCGCGGACGCGCTGCCAAGAGACTGCGAGCGCGCAGAGCGC 420
 QY 121 ArgLeuGlyAlaAspMetLysAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
 DB 421 CGGCTGGGGCGGACATGAGAGAGCTGTGCGCGCCCTGTGTACAGTACCGCGGAGAGGTG 480
 QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerLysLeuArg 160
 DB 481 CAGGCGCATGCTGGCGAGAGACACCGAGAGCTGCGGGTGGCTGCTGCCCTCCACCTGCGC 540
 QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTrp 180
 DB 541 AAGCTGCGTAAACCGGCTCTCCGCGATGCGCATGACCTGCAACAACGCTGCGCAGTGTAC 600
 QY 181 GlnAlaGlyAlaArgGluGlyAlaGlnArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 601 CAGGCGGGGGCGCGGAGAGCGCCGAGCGCGCTCAGCGCCATCGCGCAGCGCTGGGG 660
 QY 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 DB 661 CCCCTGTGTGGAACAGAGGCGCGTGGGGCGCGCCTGCTGCTGCTGCGCGCAGCGC 720
 QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGlnGluMetGly 240
 DB 721 CTACAGAGACGGGGCCAGGCTGGGGCGAGCGGCTGGCGCGCGGATGAGAGATGGCGC 780
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 DB 781 ACCCGGACCGCGGACCGCTGTGACGAGGTGAAGAGAGCAGGTGCGAGTGGCGCGCAG 840
 QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLys 280
 DB 841 CTGGAGAGACAGGCGCCAGCAGATACGCTGAGGCGGAGGCTTCCAGGCGCGCTCAAG 900
 QY 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
 DB 901 ACCTGTTGACAGCCCGCTGGTGAACATGACAGCCGCAAGTGGCGCGCTGTGTGAAGAAG 960
 QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 961 GTGACAGGCTGCGGTGGCAGCAGCGCCGCTGTGTGCCAGCAGCAATCAC 1011
 RESULT 7
 AAD22052
 ID AAD22052 standard: DNA; 1156 BP
 AC
 AAD22052;
 XX

DT 12-FEB-2002 (first entry)
 XX
 DE Human apolipoprotein E (apoE) allele, apoE2** DNA.
 XX
 KW Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;
 KW hypertriglyceridaemia; low density lipoprotein; LDL; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH CDS 61..1014
 FT /tag= a
 FT /product= "Human apoE allele, apoE2**"
 FT sig_peptide 61..114
 FT /tag= b
 FT mat_peptide 115..1011
 FT /tag= c
 FT /product= "Mature human apoE allele, apoE2**"
 XX
 PN W0200177136-A1.
 XX
 PD 18-OCT-2001.
 XX
 PF 06-APR-2001; 2001WO-US11358.
 XX
 PR 06-APR-2000; 2000US-0544386.
 PR 04-OCT-2000; 2000US-0679088.
 PR 05-APR-2001; 2001US-0827854.
 XX
 PA (KOSP-) KOS PHARM INC.
 PA (UYBO-) UNIV BOSTON.
 XX
 PI Zannis VI, Kypros KE;
 DR WPI; 2002-010885/01.
 DR P-PSDB; AAE13298.
 XX
 PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
 PT cholesterol, delaying the onset of or treating atherosclerosis in
 PT mammal, without inducing hypertriglyceridaemia
 XX
 PS Claim 14; Page 83; 91pp: English.
 XX
 CC The present sequence is a human apolipoprotein E (apoE)
 CC allele, apoE2** DNA. The apoE lipoproteins are useful for
 CC lowering cholesterol, delaying the onset of atherosclerosis,
 CC treating or regressing atherosclerosis without inducing
 CC hypertriglyceridaemia, in a mammal lacking an endogenous,
 CC normally functioning apoE gene or low density lipoprotein (LDL)
 CC receptor or is at risk for developing atherosclerosis due to
 CC accumulation of lipoprotein remnants in the bloodstream or having
 CC a defect in remnant removal.
 SQ Sequence 1156 BP; 207 A; 369 C; 432 G; 148 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 6.33e-115
 Score: 1573.00 Length: 1156
 Percent Similarity: 99.68% Matches: 315
 Best Local Similarity: 99.37% Mismatches: 1
 Query Match: 98.99% Indels: 0
 DB: 24 Gaps: 0
 US-09-827-854-18 (1-317) x AAD22052 (1-1156)
 QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 61 ATGAAGGTTCTGTGGGCTGCTGTGTCACATTCTTGCGAGATGCGCAGGCAAGTG 120
 QY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 121 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTGCGCAGACAGACCGAGTGGCAGAGC 180

QY 41 GlylnArgrTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB 181 GGCAGCGCTGGAGACCTGGCTGCTTTGGATTACCTGGCTGGTGGCAGACA 240
 QY 61 LeuSerGlnGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 DB 241 CTGTCTAGCAGCGTGCAGAGAGAGCTGCTACGCTCCAGGTCACCCAGCACTAGGGCG 300
 QY 81 LeuMetAspGlnThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 DB 301 CTGATGAGCAGACCATGAAGAGAGTGAAGCCCTACAAATCGAAGTGAAGACAACACTG 360
 QY 101 ThrProValAlaGlnGlnGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 361 ACCCGGTGGGGAGAGAGACGGGGCAGCGCTGTCCAAAGAGTGCAGGGCGCAGAGCC 420
 QY 121 ArgLeuGlnAlaAspMetGluAspValLysGlnArgLeuValGlnTyrArgGlnVal 140
 DB 421 CGGCTGGGGCGGACATGAGAGAGCTGTGGCGCGCTGGTGCAGTACCGGGCGAGAGTG 480
 QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 481 CAGGCCATGCTCGGCCAGAGACCCAGAGAGCTGGGGTGGCTCCCTCCACCTGCGC 540
 QY 161 LysLeuGlnLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 DB 541 AAGCTGGCTGAGCGGCTCTCCGATGCCGATACCTGCAAGAGCCCTGGCAGTGTAC 600
 QY 181 GlnAlaGlnAlaArgGlnGlnGlnAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
 DB 601 CAGCCCGGGGGCGGCGAGGGGGCGGCGGCTCAGCGCCATCCCGAGAGCGCTGGG 660
 QY 201 ProLeuValGlnGlnGlnGlnArgValArgAlaAlaThrValGlnSerLeuAlaGlnPro 220
 DB 661 CCCCTGGTGGAAACAGGGCCCGCTGGGGCGGCCACTGTGGGCTCCCTGGCGGCGACCG 720
 QY 221 LeuGlnGlnLysArgAlaGlnAlaThrGlnGlnArgLeuArgAlaArgMetGluGlnMetGly 240
 DB 721 CTACAGAGAGGGGGCCAGGCTGTGGGGAGCGGCTGGCGCGGAGTGGAGAGATGGGC 780
 QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
 DB 781 AGCCGGAGCCCGGACCCGCTGGAGAGGTGAGAGAGAGGAGGAGGTGGCGGCCAAG 840
 QY 261 LeuGlnGlnGlnAlaGlnGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 DB 841 CTGGAGAGAGAGCGCCAGAGATAGCGCTGGAGCGCAGGCGCTTCCAGGCCCGCTCAAG 900
 QY 281 SerTrpPheGlnProLeuValGlnAspMetGlnArgGlnTrpAlaGlnLysLeuValGlnLys 300
 DB 901 AGCTGTTTCAGAGCCCTGGGTGGAGACATGACAGCGCAGTGGGGCGGTGGAGAG 960
 QY 301 ValGlnAlaAlaValGlnThrSerAlaAlaProValProSerAspAsnHis 317
 DB 961 GTGAGAGCTGCGCTGGGACACAGCGCCGCTGTGCCACGCAATCTAC 1011
 RESULT 8
 AAF84314
 ID AAF84314 standard; cDNA; 1156 BP.
 XX
 AC AAF84314:
 XX
 DT 21-JUN-2001 (first entry)
 XX
 DE Human ApoE2 coding sequence.
 XX
 KW Human; ApoE2; Alzheimer's disease; arteriosclerosis; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 61..1014
 FT /tag= a

FT /product= "Human ApoE2"
 PN JP2001017028-A.
 XX
 XX 23-JAN-2001.
 PD
 XX 28-APR-2000; 2000JP-0128919.
 PF
 XX 06-MAY-1999; 99JP-0125647.
 PR
 XX (MITU) MITSUBISHI CHEM CORP.
 PA
 XX WPI: 2001-285406/30.
 DR P-PSDB; AAB80996.
 XX
 XX New apoE humanized mammalian cell useful for screening for agents
 PT useful for treating or preventing Alzheimer's disease and
 PT arteriosclerosis -
 XX
 PS Disclosure; Page 11-12; 22pp; Japanese.
 XX
 CC The present invention relates to an ApoE humanised mammalian cell. The
 CC present sequence is the coding sequence for human ApoE2, which was used
 CC in the method of the present invention. The ApoE humanised mammalian cell
 CC can be used for screening for agents useful for treating or preventing
 CC Alzheimer's disease and arteriosclerosis.
 XX
 SQ Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;
 Alignment Scores:
 Pred. No.: 1 31e-114 Length: 1156
 Score: 1569.00 Matches: 315
 Percent Similarity: 99.37% Conservative: 0
 Best Local Similarity: 99.37% Mismatches: 2
 Query Match: 98.74% Indels: 0
 DB: Gaps: 0
 US-09-827-854-18 (1-317) x AAF84314 (1-1156)
 QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 61 ATGAGAGTTCGTGGGCTGGCTGTGCTGTGCATCTTCGTCAGAGATGCCAGGCCAAGTGC 120
 QY 21 GluGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnTrpGlnSer 40
 DB 121 GAGCAAGCGGTGAGAGACAGAGCCGAGCCGACCTCGCCAGCAGACCCGAGTGGCAGAGC 180
 QY 41 GlylnArgrTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB 181 GGCAGCGCTGGAGACCTGGCTGCTTTGGATTACCTGGCTGGTGGCAGACA 240
 QY 61 LeuSerGlnGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 DB 241 CTGTCTAGCAGCGTGCAGAGAGAGCTGCTACGCTCCAGGTCACCCAGCACTAGGGCG 300
 QY 81 LeuMetAspGlnThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 DB 301 CTGATGAGCAGACCATGAAGAGAGTGAAGGCCCTACAAATCGAAGTGAAGACAACACTG 360
 QY 101 ThrProValAlaGlnGlnGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 361 ACCCGGTGGGGAGAGAGACGGGGCAGCGCTGTCCAAAGAGCTGCAGGGCGCAGAGCC 420
 QY 121 ArgLeuGlnAlaAspMetGluAspValLysGlnArgLeuValGlnTyrArgGlnVal 140
 DB 421 CGGCTGGGGCGGACATGAGAGAGCTGTGGCGCGCTGGTGCAGTACCGGGCGAGAGTG 480
 QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 481 CAGGCCATGCTCGGCCAGAGACCCAGAGAGCTGGGGTGGCTCCCTCCACCTGCGC 540
 QY 161 LysLeuGlnLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180

Db 541 AAGCTCCGTAAGGCGCTCCCTCCGATGCCGATGACCTGCAAGAGTGGCTGGCAGTGTAC 600
 QY 181 GlnAlaGlyAlaArgGluGluValArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db 601 CAGGCGGGGGCCCGCAGAGGGCCCGAGCCGGGCTCAGGCAATCCGCGAGCGCTGGGG 660
 QY 201 ProLeuValGluGluGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
 Db 661 CCCCTGTGGAAACAGAGGCGCGCTGGCGGGCCGACAGTGGGCTCCCTGGCCGCGCAGCGG 720
 QY 221 LeuGluGluArgAlaGluAlaThrPglGlyArgLeuArgAlaArgMetGluGluMetGly 240
 Db 721 CTACAGGAGCGGGCCCGAGGCTGTGGCGCGAGCGCGCTGCGCGGATGAGAGATGGGCG 780
 QY 241 SerArgThrArgAspArgLeuAspGluValLeuGlyGluGluValAlaGluValArgAlaGly 260
 Db 781 AGCCGAGACCCGCGACCGCTGGACGAGGTGAAGAGACAGGTGGAGGTGGCGCGCAAG 840
 QY 261 LeuGluGluGluAlaGluGluGluIleArgLeuGluAlaGluAlaPheGluAlaArgLeuGly 280
 Db 841 CTGGAGAGACAGGCGCCAGAGATAGCCCTGCAAGCGCGAGCCCTTCAGGCGCGCTCAAG 900
 QY 281 SerTrpPheGluProLeuValGluAspMetGluArgGluTrpAlaGlyLeuValGluGly 300
 Db 901 ACCTGTGTGAGCCCTGTGTGAGAGACATGACAGCGCCAGTGGGCGCGGTGGAGAG 960
 QY 301 ValGluAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 961 GTGCAAGGCTGCGGTGGGACACAGCGCCGCTGTGCTCCAGCGACATCAC 1011
 RESULT 9
 AAD22049
 ID AAD22049 standard; DNA; 1156 BP.
 AC AAD22049;
 DT 12-FEB-2002 (first entry)
 DE Human apolipoprotein E (apoE) isoprotein, apoE2 DNA.
 XX
 KW Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;
 KM hypertriglyceridaemia; low density lipoprotein; LDL; ds.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 61..1014
 FT sig_peptide /tag= a
 FT mat_peptide /tag= b
 FT /tag= C
 FT /Product= "Mature human apoE isoprotein, apoE2"
 XX
 PN WO200177136-A1.
 PD 18-OCT-2001.
 XX
 PF 06-APR-2001; 2001WO-US11358.
 XX
 PR 06-APR-2000; 2000US-0544386.
 PR 04-OCT-2000; 2000US-0679088.
 PR 05-APR-2001; 2001US-0827854.
 XX
 PA (KOSP-) KOS PHARM INC.
 PA (UYBO-) UNITV BOSTON.
 XX
 PI Zannis VI, Kypreos KE;
 XX
 DR WPI: 2002-010885/01.
 DR P-PSDB: AAE13295.
 XX

PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
 PT cholesterol, delaying the onset of or treating atherosclerosis in
 PT mammal, without inducing hypertriglyceridaemia -
 PS Claim 14; Page 81-82; 91pp; English.
 XX
 CC The present sequence is a human apolipoprotein E (apoE)
 CC isoprotein, apoE2 DNA. The apoE lipoproteins are useful for
 CC lowering cholesterol, delaying the onset of atherosclerosis,
 CC treating or regressing atherosclerosis without inducing
 CC hypertriglyceridaemia, in a mammal lacking an endogenous,
 CC normally functioning apoE gene or low density lipoprotein (LDL)
 CC receptor or is at risk for developing atherosclerosis due to
 CC accumulation of lipoprotein remnants in the bloodstream or having
 CC a defect in remnant removal.
 XX
 SQ Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;
 Alignment Scores:
 Pred. No.: 1..31e-114 Length: 1156
 Score: 1569.00 Matches: 315
 Percent Similarity: 99.37% Conservative: 0
 Best Local Similarity: 99.37% Mismatches: 2
 Query Match: 98.74% Indels: 0
 DB: 24 Gaps: 0
 US-09-827-854-18 (1-317) x AAD22049 (1-1156)
 QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGluAlaVal 20
 Db 61 ATGAAGGTTCTGTGGGCTCGTTGGTGCATCTTCGAGATGCCAGGCCAAGCTG 120
 QY 21 GluGluAlaValGluThrGluProGluProGluLeuArgGluGluThrGluTrpGlnSer 40
 Db 121 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCCACAGACCGAGTGGCAGAC 180
 QY 41 GlyGluArgTrpGluLeuAlaLeuGluArgPheTrpAspTrpLeuArgTrpValGluThr 60
 Db 181 GGCACAGCGCTGGAGACTGCGACTGGGTGCTTTGGGATTAAGTGGGTGGAGAGA 240
 QY 61 LeuSerGluGluValGluGluGluLeuLeuSerSerGluValThrGluLeuArgAla 80
 Db 241 CTGTCTGACAGAGTGCAGAGAGAGCTGCAGCTCCAGCTCACCGAGACTGAGGCGG 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGluGlu 100
 Db 301 CTGATGACGAGACCACTGAAGAGTTGAAGGCTTCAAAATCGAAGCTGAGGACACTG 360
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGluAlaAlaGluAla 120
 Db 361 ACCCGGTGGCGGAGAGAGCGGGCGAGCGTGTCCAAGAGACTGCAAGCGCGCAGGCG 420
 QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGluTrpArgGlyGluVal 140
 Db 421 CGGCTGGGCGCGACATGAGAGACGTGCGCGCGCTGTGCAGTACCGCGCGGAGGTG 480
 QY 141 GlnAlaMetLeuGluGluSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 481 CAGGCCATGCTGGCGCAGAGCACCGAGAGCTGCGGGTGGCTGCGCTCCACCTGGCGG 540
 QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGluLysArgLeuAlaValTrp 180
 Db 541 AAGCTGCGTAAGCGGCTCTCGCGCATGGCGATGACCTCAGAAAGTCCCTGGCAGGTAC 600
 QY 181 GlnAlaGlyAlaArgGluGluValArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db 601 CAGGCGGGGGCCCGCAGAGGGCCCGAGCGCGCTCAGGCAATCCGCGAGCGCTGGGG 660
 QY 201 ProLeuValGluGluGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
 Db 661 CCCCTGTGGAAACAGAGGCGCGCTGGCGGGCCGACAGTGGGCTCCCTGGCCGCGCAGCGG 720
 QY 221 LeuGluGluArgAlaGluAlaThrPglGlyArgLeuArgAlaArgMetGluGluMetGly 240

Db	721	CTACAGAGGAGCGGGCCACGAGCCCTGGGGGCGAGCGGCTGGCGCGCGCGCGAGATGGAGAGATGGC	780
Qy	241	SerArGthrArGAsPArGLeuAsPGLuValLysGLuGLuValAlaValAlaValAlaValLys	260
Db	781	AGCCGAGACCCCGACGCCCTCGAGAGAGGTGAAGAGCAGAGGTGGCGGAGGTGGCGCCAAAG	840
Qy	261	LeuGLuGLuGLuAlaGLuGLuIleIaArgLeuGLuIlaGLuAlaPheGLuIlaIaArgLeuLys	280
Db	841	CTGGAGAGGAGCAGGGCCACACAGATACGCGCTGCAGGCGCAGAGCGCTTCCAGGCGCCCTCAAG	900
Qy	281	SerTrpPheGLuPProLeuValGLuAsPmectIaArgGLuInTrpAlaGLuLeuValGLuLys	300
Db	901	AGCTGGTTCCGAGCCCTTGTTGTGAAAGACATGACAGGCCACAGTGGCGCGGCTGGTGGAGAG	960
Qy	301	ValGLuIlaValAlaValGLuThrSerAlaAlaProValProSerAspAsnHis	317
Db	961	GTCGAGGCGTGGCGTGGGGCAGCGCGCGCCCTGTGGCCAGCAGCATTCAC	1011
RESULT 10			
AAAT06957	ID	AAAT06957 standard; cDNA to mRNA; 1157 BP.	
XX	AC	AAAT06957;	
XX	AC		
DT	DT	19-JUN-1996 (first entry)	
XX	XX		
DE	DE	Human apolipoprotein-E (ApoE) cDNA.	
XX	XX		
KW	KW	Recombinant; human; apolipoprotein-E; ApoE; insect cells; larva;	
KW	KW	Manduca sexta; Autographa californica nuclear polyhedrosis virus;	
XX	XX	haemolymph; lipid complex; biologically active; ds.	
OS	OS	Homo sapiens.	
XX	XX		
Key	Key	Location/Qualifiers	
FT	FT	misc_feature 62..119	
FT	FT	/*tag= a	
FT	FT	note= "misc_signal"	
FT	FT	62..1015	
FT	FT	/*tag= b	
XX	XX		
PM	PM	US5472858-A.	
XX	XX		
PD	PD	05-DEC-1995.	
XX	XX		
PF	PF	04-JUN-1991; 91US-0709949.	
XX	XX		
PR	PR	04-JUN-1991; 91US-0709949.	
XX	XX		
PA	PA	(WISC) WISCONSIN ALUMNI RES FOUND.	
XX	XX		
PI	PI	Attie AD, Beckage NE, Gretch DG, Sturley SL;	
XX	XX		
DR	DR	WPI; 1996-029812/03.	
XX	XX	P-PSDB; AAR86791.	
PT	PT	Prodn. of recombinant apo:lipoprotein E in insects - by infecting	
PT	PT	Manduca sexta larvae with recombinant Autographa californica	
PT	PT	nuclear polyhedrosis baculovirus vector.	
XX	XX		
XX	XX		
XX	XX		
CC	CC	Disclosure; Columns 11-14; 10pp; English.	
CC	CC	Recombinant human apolipoprotein-E (ApoE) (AAR86791) can be produced	
CC	CC	by preparing a genetic construct (contg. an ApoE-encoding sequence,	
CC	CC	e.g. AAT06957, and flanking regulatory sequences enabling the protein	
CC	CC	to be expressed in insect cells), which is then introduced into a	
CC	CC	Manduca sexta larva (using a recombinant Autographa californica	
CC	CC	nuclear polyhedrosis virus) and recovering the protein from the	
CC	CC	haemolymph of the larval host. The ApoE produced is in a form	
CC	CC	sufficiently complexed with lipids to be biologically active, which	
CC	CC	cannot be achieved in insect cell cultures, and can therefore be	
CC	CC	used in therapeutic applications	

XX	Sequence	1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;
Alignment Scores:		
Pred. No.:	1,57e-114	Length: 1157
Score:	1568.00	Matches: 314
Percent Similarity:	99.05%	Conservative: 0
Best Local Similarity:	99.05%	Mismatches: 3
Query Match:	98.68%	Indels: 0
DB:	17	Gaps: 0
US-09-827-854-18 (1-317) x AAT06957 (1-1157)		
QY	1 MetLysValLeuTPpAlaLeuLeuValThrpHeuAlaGlyGlnAlaLysVal	20
DB	62 ATGAAGGTTCTGTGGCGCTGCTCTGTCATCTTCGACGAGATGCCAGGCCAAAGGTG	121
QY	21 GlnGlnAlaValAlaGluThrGluProGluProLeuLeuArgGlnGlnThrGluTrpGlnSer	40
DB	122 GAGCAAGCGGCTGGAGACAGACGCCGAGCCGACGCTGGCCAGCAGACGAGCTGGCAGAC	181
QY	41 GlyIleArgTrpGluLeuAlaLeuGlyYArpHeTrpAspTrpLeuArgTrpValGlnThr	60
DB	182 GGCCAGCCCTGGGAACCTGGCACTGGGTGCTTTGGATTACCTGGCCCTGGCTGCACACA	241
QY	61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrpGlnGluLeuArgAla	80
DB	242 CTGTCTGAGCAGGTCCAGAGAGACCTGCTGCATGCCAACGACCCAAAGACTGAGGGCG	301
QY	81 LeuMeTAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu	100
DB	302 CTGATGGAGCAGACCATTAAGAGATTGAAGGCTTCAACAATGGAACTGGAGAAACAATG	361
QY	101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
DB	362 ACCCGGTAGCGGAGAGAACCCGGGCAACGGCTGTCTCAAGAGAGCTGACAGCGCGCAGGCC	421
QY	121 ArgLeuGluYAlaAspMetGluAspValCysGluYArgLeuValGlnTrpArgGluVal	140
DB	422 CGGCTGGGCGCGACATGAGAGACCTGTGGCGCCCTGTGTGCATGACCGCGGAGGTG	481
QY	141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg	160
DB	482 CAGGCATGCTCTGGCGACAGACACGAGAGACTGTGGGTGCGCTCGCCCTCCACCTCGC	541
QY	161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
DB	542 AAGGTCCTTAAGCGCGCTCTCTCCGGATCCCGATACCTTGCAAGAACCCCTGGCAGTGTAC	601
QY	181 GlnAlaGluYAlaArgGluGluYAlaGluArgGlyLeuSerAlaIleArgGluAlaArgLeuGly	200
DB	602 CAGGCCGGGGCGCCGAGAGGCGCCGAGCGCGCTTCAGCGCATCCGCGAGCGCTGGGG	661
QY	201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro	220
DB	662 CCCCTGGTGAACAGGGCGCGCTCGGGCGCCACTGTGGGCTCTCCGTGGCGGACAGCCG	721
QY	221 LeuGlnGluValArgAlaGlnAlaTrpGlyCysLysArgLeuArgAlaArgMetGluGluMetGly	240
DB	722 CTACAGGAGCGGGCCCAAGGCTTGGGGGAGAGGCTGTGGCGCGCGCGAGTGGAGAGATGGG	781
QY	241 SerArgTrpArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys	260
DB	782 AGTCGAGACCCGACCGCGCTTGACAGAGTGAAGAGACAGGTGGCGGAGTGGCCGCAAG	841
QY	261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys	280
DB	842 CTGGAGGAGCGAGGCCACAGCAATCGCTCGAGGCGCAAGGCTTTCAGGCGCCGCTCAAG	901
QY	281 SerTrpPheGluProLeuValGluAspMetGlnArgGluInTrpArgLysLeuValGluLys	300
DB	902 AGCTGGTTCGAGCCCTGTGGTGAAGACATGACGCGCAAGTGGCGGGGCTGTGTGGAAG	961

XX WO200212440-A2.
 XX 14-FEB-2002.
 XX 07-AUG-2001; 2001WO-US24708.
 XX 07-AUG-2000; 2000US-223323P.
 XX 05-JUN-2001; 2001US-0873319.
 XX (GENE-) GENE LOGIC INC.
 PA (NIBS) JAPAN TOBACCO INC.
 PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
 DR WPI; 2002-257476/30.
 XX
 PT Identifying drugs for and diagnosing benign prostatic hyperplasia, by
 PT detecting expression levels of one or more genes in prostate cells from
 PT patient that are differentially regulated compared to normal prostate
 PT cells -
 XX
 PS Disclosure; Page 239-240; 444pp; English.
 CC The invention relates to a method of diagnosing (I) the onset or
 CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
 CC or identifying an agent that modulates the onset or progression of BPH.
 CC The method is based on changes in gene expression in BPH tissue isolated
 CC from patients exhibiting different clinical states of prostate
 CC hyperplasia as compared to normal prostate tissue. (I) comprises
 CC detecting the expression levels of one or more genes in prostate cells
 CC from the subject that are differentially regulated compared to normal
 CC prostate cells. (II) comprises preparing a first gene expression profile
 CC of BPH cells or BPH-like cell population, exposing the cells to the
 CC agent, preparing a second gene expression profile of the agent exposed
 CC cells, and comparing the first and second gene expression profiles.
 CC (I) is useful for diagnosing the onset or progression of BPH. (II) is
 CC useful for identifying an agent that modulates the onset or progression
 CC of BPH. The methods are useful to present information identifying
 CC the expression level in a tissue or cells, by comparing the expression
 CC level of genes given in the specification in the tissue or cells to the
 CC level of expression of gene in the database, and displaying the
 CC expression levels of at least one gene in the tissue or cell sample
 CC compared to the expression level in BPH. Agents using (II) are useful for
 CC treating BPH or prostate cancer. ABR64106-ABR64860 represent human
 CC benign prostatic hyperplasia gene sequences of the invention.
 CC
 XX
 SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.57e-114 Length: 1157
 Score: 1568.00 Matches: 314
 Percent Similarity: 99.05% Conservative: 0
 Best Local Similarity: 99.05% Mismatches: 3
 Query Match: 98.68% Indels: 0
 DB: 24 Gaps: 0
 US-09-827-854-18 (1-317) x ABR64514 (1-1157)
 OY 1 MethylValleuTrpAlaAlaLeuLeuValThrpheLeuAlaGlyCysGlnAlaVal 20
 DB 62 ATGAAGGTCTGTGGGTGGCTGTCTGTCACTCTGGAGAGATCCAGGCCAAGGTG 121
 OY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 122 GAGCAAGCGGTGGAGAGACAGCGGAGCCGAGCTGGCCAGCAGACCGAGTGCACAGC 181
 OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTTPValGlnThr 60
 DB 182 GGCCAGGCGTGGGAACTGGCACTGGCTTTGGGATTACCGCGCTGGGTGCACAGCA 241
 OY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerGlnValThrpGlnGluLeuArgAla 80
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 242 CTGTCTAGCAGAGTGCAGAGAGACCTGTCTAGCTCCCAAGTCCACCAAGACTGAGGCG 301
 OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 DB 302 CTGATGGACGAGACCATGAAAGAGTTGAAGCCCTACAAATCGAACTGGAGAACTG 361
 OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
 DB 362 ACCCGGTAGCGAGAGACCGCGGACAGCGCTGTCCAAAGAGCTGACAGCGCCAGGCG 421
 OY 121 ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 DB 422 CGGCTGGGGCGGACATGAGAGACGTGTGGCGCCGCTGTGTACATCCCGGGAGAGTG 481
 OY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 482 CAGGCCATGCTCGGSCCAGAGACAGCAGAGAGTGGGGTGGCCCTCGGCTCCACCTGGC 541
 OY 161 LysLeuGlyLysArgLeuLeuArgAlaAspAlaAspLeuGlnLysArgLeuAlaValTyr 180
 DB 542 AAGCTGGTAAAGCGGCTCTCCGCGATCCCGATGACCTGCAGAAAGCCCTGGCAGTGTAC 601
 OY 181 GlnAlaGlyAlaArgGluGluGluAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 602 CAGCGCGGGGCGCCGAGGGCGCCGAGCGGCTCAGCGCCATCCCGAGCGCTGGGG 661
 OY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 DB 662 CCCCTGGTGAACAGGGGCCGCTGGCGCCGACCTGTGGCTCTCCCTGGCGGCGACCG 721
 OY 221 LeuGlnGluArgAlaGlnAlaATTPGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
 DB 722 CTACAGAGAGCGGGCCGAGCGCTGGCGGACGCGCTGGCGCCCGCCGAGTGCAGATGGCC 781
 OY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
 DB 782 AGTGGACCCGCGACCCCTGGAGAGAGTGAAGAGCAGAGTGGCGAGGTGGCGCCAG 841
 OY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 DB 842 CTGGAGAGAGCAGGCGCCAGAGATAGCGCTGCAGGCGCCAGCGCTTCAGCGCCGCTCAAG 901
 OY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 DB 902 AGCTGGTTCAGACCCCTGGGTGAAGACATCAGCGCCAGTGGCGGCTGTGTGAGAGAG 961
 OY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 962 GTGCAGGCTCGCTGGGACACAGCGCCGCTGTGCCACGACCAATTCAC 1012
 RESULT 13
 ABL65450
 ID ABL65450 standard; DNA; 1157 BP.
 XX ABL65450;
 XX 15-MAY-2002 (first entry)
 DE Lung cancer related gene sequence SEQ ID NO:3787.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; aniltumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US10838.
 XX

PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 26-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX (AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI: 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1: SEQ ID 3787; 44pp: English.
XX
XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (1) of a signature gene set, where (1)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61654
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (1) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.

XX
SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;
Alignment Scores:
Pred. No.: 1,57e-114 Length: 1157
Score: 1568.00 Matches: 314
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 99.05% Mismatches: 3
Query Match: 98.68% Indels: 0
DB: 24 Gaps: 0
US-09-827-854-18 (1-317) x ABL65450 (1-1157)
QY 1 MetLysValLeuThrPalaIalaLeuValThrPheLeuAlaGlyCysGlnAlaTyrVal 20
DB 62 ATGAAGTTCTGTGGGCTCGTTGCTGTCACATTCCTGCGAGATGCCAGCAAGGTG 121
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 122 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTCGCCACAGACAGACGATGGCAGAGC 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnTrp 60
DB 182 GGCACAGCGTGGGAAGTGGACTGGGTCCCTTTGGATTACCTGCGTGGGTGCAGACA 241
QY 61 LeuSerGluGlnValGlnGluLeuLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 242 CTGTCTGACAGAGTGCAGAGAGAGCTGCTCAGCTCCCAAGTCAACCAAGACTGAGAGCGG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGluLeu 100
DB 302 CTGATGACGAGACCATGTAAGAGTTGAAGGCTTACAAATCGAAGTGAAGCAACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
DB 362 ACCCGGTAGCGAGAGAGAGCGGGGCGACGCTCTCCAAAGAGCTGCAGAGCGCGAGGCC 421
QY 121 ArgLeuGluAlaAspMetLysAspValLysGlyArgGluValGlnThrArgGlyGluVal 140
DB 422 CGGCTGGGCGGACATGAGAGAGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
QY 141 GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 482 CAGGCCATGCTCGGCCAGAGCAACCGAGAGCTGGGGTGGCTGGCTGCCCACTGGCGGC 541
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 542 AAGCTGCGTAAAGCGGCTCCCGCGATCCCGATGACCTGCAGAAAGCGCTGGCAGGTAC 601
QY 181 GlnAlaGluAlaArgGluGluValArgGlyLysLeuSerAlaIleArgGluArgLeuGly 200
DB 602 CAGCGCGGGGCGGCGAGGGGCGCGAGCGGCGGCTCAGCGCCATCCGAGCGGCTGGGG 661
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 662 CCCGTGGTGAACAGAGGCGCGTGGCGGCGGCACATGTGGGCTCCCTGGCCGCGCAGCG 721
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
DB 722 CTACAGGAGCGGCGCGAGCGCTGGGGCGAGCGGCTGCGGCGCGATGAGAGATGGGCG 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaTyr 260
DB 782 ACTCGGACCGCGGACCGCTCGAGACAGGTGAAGAGAGAGTGGCGAGGTGGCGCGCAAG 841
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
DB 842 CTGAGGAGAGCAGGCGCAGAGATACGCTCGAGGCGGAGAGCTTCCAGGCGCGCTCAAG 901
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 902 ACCTGTTTCGAGCCCTGTGTGAAGACATGACAGCGCCATGTGGCGGCTGTGTGAAGAG 961

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:17:52 ; Search time 34.7343 Seconds
(without alignments)
2798.866 Million cell updates/sec

Title: US-09-827-854-18

Sequence: 1 MKVIMALLVTLACGCAQV.....VEKVOAAVTSAPVPSDNH 317

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Command line parameters:
-MODEL=frimex.p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09827854/rnucat.11032003.101611.27510/app.query.fasta.1.3576
-DB=Issued Patents.NA -OPMT=fastap -SUPER=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09827854.@CGL1.1.224.@rnucat.11032003.101611.27510 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEGLScores=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-NARN.TIMEOUT=30 -TREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued Patents.NA.*

1: /cgn2.6/prodata/2/ina/5A.COMB.seq:*
2: /cgn2.6/prodata/2/ina/5B.COMB.seq:*
3: /cgn2.6/prodata/2/ina/5A.COMB.seq:*
4: /cgn2.6/prodata/2/ina/5B.COMB.seq:*
5: /cgn2.6/prodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2.6/prodata/2/ina/Backfilest1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1568	98.7	1157	1	US-07-709-949-1
2	1130	71.1	1126	4	US-08-949-155-5
3	1130	71.1	1126	4	US-09-819-964-5
4	994	62.6	4267	4	US-08-949-155-51
5	994	62.6	4267	4	US-09-819-964-51
6	976	61.4	660	2	US-08-726-306A-28
7	503	31.7	330	1	US-07-849-389-6
8	366	23.0	252	3	US-08-617-256-24
9	366	23.0	252	4	US-09-287-141-24
10	366	23.0	252	4	US-09-431-613-24
11	366	23.0	252	4	US-09-504-245-24
12	366	23.0	252	4	US-09-287-682-24

13	366	23.0	252	4	US-09-287-679-24	Sequence 24, Appl
14	366	23.0	252	4	US-09-397-766-24	Sequence 24, Appl
15	366	23.0	252	4	US-09-287-681-24	Sequence 24, Appl
16	366	23.0	252	4	US-09-495-444-24	Sequence 24, Appl
17	177.5	11.2	842	4	US-08-952-796-1	Sequence 1, Appl
18	174.5	11.0	801	1	US-07-959-946-4	Sequence 4, Appl
19	174.5	11.0	801	1	US-08-333-577-4	Sequence 4, Appl
20	174.5	11.0	801	5	PCT-US92-08634-4	Sequence 2, Appl
21	168	10.6	5661	4	US-08-938-105-2	Sequence 4, Appl
22	157.5	9.9	964	1	US-08-448-606-5	Sequence 5, Appl
23	154.5	9.7	863	1	US-08-448-606-7	Sequence 7, Appl
24	153	9.6	1879	4	US-09-750-580-2	Sequence 2, Appl
25	150	9.4	3256	2	US-08-968-751-3	Sequence 3, Appl
26	147.5	9.3	603	4	US-08-952-796-14	Sequence 14, Appl
27	147	9.3	8789	1	US-08-328-254-5	Sequence 2, Appl
28	147	9.3	10136	1	US-08-353-700-2	Sequence 2, Appl
29	147	9.3	10136	5	PCT-US95-16216-2	Sequence 12, Appl
30	146	9.2	4868	1	US-08-139-937-12	Sequence 12, Appl
31	146	9.2	4868	5	PCT-US93-11310-12	Sequence 12, Appl
32	140.5	8.8	1960	2	US-08-533-306A-1	Sequence 1, Appl
33	140.5	8.8	1960	2	US-08-742-923A-1	Sequence 1, Appl
34	140.5	8.8	11236	1	US-07-853-913-1	Sequence 1, Appl
35	139.5	8.8	4852	1	US-07-853-913-3	Sequence 3, Appl
36	139	8.7	516	4	US-09-183-861-34	Sequence 34, Appl
37	139	8.7	516	4	US-09-022-765-34	Sequence 34, Appl
38	137	8.6	6706	1	US-08-195-487-3	Sequence 3, Appl
39	137	8.6	6706	5	PCT-US93-06160-3	Sequence 3, Appl
40	136.5	8.6	3319	1	US-08-006-676B-2	Sequence 2, Appl
41	136.5	8.6	3319	1	US-08-282-845-1	Sequence 1, Appl
42	136.5	8.6	3319	2	US-08-428-414A-4	Sequence 4, Appl
43	136.5	8.6	3319	5	PCT-US94-00324-2	Sequence 2, Appl
44	136.5	8.6	6306	1	US-08-466-390-3	Sequence 3, Appl
45	136.5	8.6	6306	1	US-08-470-950-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-07-709-949-1
; Sequence 1, Application US/07709949
; Patent No. 5472858
; GENERAL INFORMATION:
; APPLICANT: Attie, Alan D
; APPLICANT: Gretch, Daniel G
; APPLICANT: Sturley, Stephen L
; APPLICANT: Beckage, Nancy E
; TITLE OF INVENTION: Production of Recombinant Proteins in
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: P.O. Box 2113
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/709,949
; FILING DATE: 19910604
; CLASSIFICATION: 433
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 9629691801
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 1:

QY	1	MettysValLeuTPPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaVal	20
Db	62	ATGAAGGTTCTGTGGCGCTGGTCTGGTCAATTCTGGCAGGATGCCAGCCAAAGTG	122
QY	21	GIuGlInAlaValAlGIuThrGluProGluProGluLeuArgGlnInThrGIuTPGInSer	40
Db	122	GAGCAAGCGGGGAGACAGACCCCGAGCTGGCCAGCGACGACGAGTGGCAAGC	182
QY	41	GIyGlnAArgTPRGIuLeuAlaLeuGlyAARpHeTPAspTRYLeuAArgTPRValGInThr	60
Db	182	GGCCAGCGCTGGGAACGTGGCACTGGTGCGTTTGGGATTACCTGGCGCTGGGTGCACACA	242
QY	61	LeuSerGIuGlnValGInGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80
Db	242	CTGTCTGACGAGGGTCAGGAGAGACTCTCACTGCTCCAAAGTACCCAAAGCAATGAGGGCG	302
QY	81	LeuMetAspGIuThrMetLeuGlyLeuLeuGlyAlaTYrLeuSerGIuLeuGluGluLeu	100
Db	302	CTGATGGACGACCATGAAGAGATTGAAGCCCTCAAAATGGAACTGGAGAACCACTG	362
QY	101	ThrProValAlaGluGluThrArgAlaArgLeuSerLeuGlnAlaAlaGlnAla	120
Db	362	ACCCCGAGGCGGAGGAGACCGGGCCAGCTGTCCAAAGAGCTGCGACGCGCCAGGCC	422
QY	121	ArgLeuGlnAlaAspMetGIuAspValCysGlyArgLeuValGlnTYrArgGlyGluVal	140
Db	422	CGCGTGGCGCGGCATGTGAAGACGTGTGGGGCGCGCTGGTGACGTACCGGGGAGGTG	482
QY	141	GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	482	CAGGCCATGCTCGGCCAGACACCGAGAGACTGGGGGTGGCCCTGCCCTCCACCTGGCG	542
QY	161	LysLeuAspLysArgLeuLeuArgAspAlaAspAspLeuGlnTyrArgLeuAlaValTYr	180
Db	542	AAGCTCGCTAAGCGGCGTCTCTCCGCGATCCCGATACCTGCAGAAAGCCCTGGCAGTGTAC	602
QY	181	GlnAlaGluValaArgGluGluGlyAlaGluArgGlyLeuSerAlaIleArgGluIuArgLeuGly	200

[illegible]

RESULT 2

```

Sequence 5, Application US/08949155
Patent No. 6271436
GENERAL INFORMATION:
APPLICANT: Piedrahita, Jorge A
TITLE OF INVENTION: Compositions and Methods for the
TITLE OF INVENTION: Generation of Transgenic Animal Species
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE AND DURKEE
STREET: P. O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/949,155
FILING DATE: Concurrently Herewith
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/027,338
FILING DATE: 11-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/046,094
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hibler, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TAM:177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:

```


NAME/KEY: CDS
LOCATION: 51..1001
US-08-949-155-5

Alignment Scores:
Pred. No.: 1,65e-102
Score: 1130.00
Percent Similarity: 83.44%
Best Local Similarity: 69.69%
Query Match: 71.11%
DB: 4

Length: 1126
Matches: 223
Conservative: 44
Mismatches: 45
Indels: 8
Gaps: 3

US-09-827-854-18 (1-317) x US-08-949-155-5 (1-1126)

QY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 51 ATGAGGCTTCTGTGGTCTTGTGGTACCTCTCGACAGATGCCGACAGAGAC 110
QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
DB 111 GAGCCGGGG-----CCGCCCGGAGGTGCACGTGTGTGGAGAGAGCCCAAG 158
QY 38 TrpGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrp 57
DB 159 TGGCAGGCGACGACGCCCTGGGAGCGCCCTGCGCTCTGTGGATTACCTGCGCTGG 218
QY 58 ValGlnThrLeuSerGluGlnAlaGlnGluGluLeuLeuSerSerGlnValThrGlnGlu 77
DB 219 GTGCACTCCCTCTGTACCAAGGCGAGGAGAGCTGCTGACACCAAGGTACCCACGAA 278
QY 78 LeuArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlu 97
DB 279 CTGACGAGACTGTATAGAGAGACATGAGAGGTGAAGCTACCGCCAGAGAGCTGGAG 338
QY 98 GluGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAla 117
DB 339 GCGCACTCGGGCCCTGTACCGAGGAGCGCGCCCTGTCCAAAGAGCTGACGAGCG 398
QY 118 AlaGlnAlaArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArg 137
DB 399 GCGCAGGCGCGGTGGGCGCCGACATGAGAGAGCTGCCAACCGCTGTGTCTACCCG 458
QY 138 GlyGluValGlnAlaMetLeuGlnLysSerThrGluLeuArgValArgLeuAlaSer 157
DB 459 AGCAGAGTGCACACATGTGTGGGCCAGACACCGAGAGACTGCGAGCGCCCTGCTTC 518
QY 158 HisLeuArgLysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeu 177
DB 519 CACCTCGCAGACTGCGCAGGCGCTGCCGACACCGAGACTGCGAAGAGCCCTG 578
QY 178 AlaValTyrGlnAlaGlyAlaArgGluAlaArgGlyLysSerAlaLeuArgGlu 197
DB 579 GCGGTGTACAGGCGGGCTGCGCGAGGCGCGAGCGAGCTGAGTGCCTCCGCGAG 638
QY 198 ArgLeuGlyProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAla 217
DB 639 CGCCTCGGGCCCTGTGTGAGCAGGCGCATTTGCGCGCCGCCACCTGAGTACAGAGCC 698
QY 218 GlyGlnProLeuGlnGluArgAlaGlnAlaTyrGlyGluArgLeuArgAlaArgMetGlu 237
DB 699 GGCAGGCCCTCGCGAGCGCGCGAGGAGCTGCGCGCAAGACTGCGCGAGCGCTGAG 758
QY 238 GluMetGlySerArgTrpArgAspArgLeuAspGluValLysGluGlnAlaAlaGluVal 257
DB 759 GAGATGGGAGCGCGAGCGCGCGAGCTGAGATGAGATCGTGAAGCTGAGAGCGTGG 818
QY 258 ArgAlaLysLeuGluGlnGluAlaGlnGlnLeuArgGlnAlaGluAlaLysVal 277
DB 819 CCGACCAAGGTGAGAGAGAGGAGCGACCGCTGCGCTCAGGCCAGGAGATTCCAGCGC 878
QY 278 ArgLeuLysSerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeu 297
DB 879 CTCCTCAAGGCTGTTGAGCGCTGCGTGAAGACATACGCGCGCAGTGGCGCGGCTG 938

QY 298 ValGluLysValGlnAlaValAlaGly---ThrSerAlaAlaProValProSerAspAsn 316
DB 939 GTGGAGAGATGATGAGCGCGTGCAGATATAGCTCTCCACCTTGCAGCTGATATAT 998

RESULT 3
US-09-819-964-5
Sequence 5, Application US/09819964
Patent No. 6369294
GENERAL INFORMATION:
Applicant: Piedrahita, Jorge A
Bazer, Fuller W
TITLE OF INVENTION: Compositions and Methods for the
Generation of Transgenic Animal Species
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE AND DURKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/819,964
FILING DATE: 28-Mar-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/949,155
FILING DATE: <unknown>
APPLICATION NUMBER: US 60/046,094
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hibler, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TANK:177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 51..1001
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-819-964-5

Alignment Scores:
Pred. No.: 1,65e-102
Score: 1130.00
Percent Similarity: 83.44%
Best Local Similarity: 69.69%
Query Match: 71.11%
DB: 4

Length: 1126
Matches: 223
Conservative: 44
Mismatches: 45
Indels: 8
Gaps: 3

US-09-827-854-18 (1-317) x US-09-819-964-5 (1-1126)

QY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 51 ATGAGGCTTCTGTGGTCTTGTGGTACCTCTCGACAGATGCCGACAGAGAC 110
QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
DB 111 GAGCCGGGG-----CCGCCCGGAGGTGCACGTGTGTGGAGAGAGCCCAAG 158


```

QY 38 TrpGlnSerGlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrp 57
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 159 TGGCAGGAGCAGCCAGCCCTGGAGCAGAGCCCTGGCGCTTCTGGGATTAACCTCGCTGG 218
QY 58 ValGlnThrLeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlu 77
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 219 GTGAGAGCCCTGCTGACCAAGTGCAGAGAGAGCTGCTGACCAAGAGTCCAGAGAA 278
QY 78 LeuArgAlaLeuMetAspGlnThrMetGlyGlnLeuLysAlaTyrLysSerGlnLeuGlu 97
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 279 CTGACGAGAGCTGATAGAGAGAGATGAGAGAGTGAAGGCTTACCCGAGAGACTGGAG 338
QY 98 GluGlnLeuThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAla 117
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 339 GCGAGAGCTGGGCCCTGGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 398
QY 118 AlaglnAlaArgLeuGlnValAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArg 137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 399 GCGAGAGCCGCGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 458
QY 138 GlyGlnValGlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSer 157
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 459 AGCAGAGTGCACACATGTTGGGCCAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 518
QY 158 HisLeuArgLysLeuGlyLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeu 177
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 519 CACCTGGCAGAGCTGCGCAAGCGGCTCTCCGACACAGAGAGAGAGAGAGAGAGAGAGAG 578
QY 178 AlaValThrGlnAlaGlnValArgGlnGlnValAlaGlnValArgGlnLysSerAlaIleArgGlu 197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 579 GCCGTGACCAAGCGGGGCTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 638
QY 198 ArgLeuGlnProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAla 217
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 639 CGCCTCGGCGCCCTGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 698
QY 218 GlyGlnProLeuGlnGlnValArgAlaGlnAlaThrGlnGlnArgLeuArgAlaArgMetGlu 237
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 699 GCGCAGAGCGCTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 758
QY 238 GluMetGlySerArgTrpArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnVal 257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 759 GAGATGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 818
QY 258 ArgAlaLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 277
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 819 CGCACAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 878
QY 278 ArgLeuLysSerTrpPheGlnProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeu 297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 879 CTCTCAAGAGCTGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 938
QY 298 ValGlnThrLeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlu 316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 939 GTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 998

```

```

RESULT 4
US-08-949-155-51
; Sequence 51, Application US/08949155
; Patent No. 6271436
GENERAL INFORMATION:
; APPLICANT: Pledranta, Jorge A
; APPLICANT: Bazer, Fuller W
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Generation of Transgenic Animal Species
; NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE AND DURKEE
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: US
; ZIP: 77210-4433

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/949,155
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/027,338
; FILING DATE: 11-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,094
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hidler, David W.
; REGISTRATION NUMBER: 41,071
; REFERENCE/DOCKET NUMBER: TAMR:177
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4267 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-949-155-51

Alignment Scores:
Pred. No.: 2.53e-88 Length: 4267
Score: 994.00 Matches: 217
Percent Similarity: 58.20% Conservative: 42
Best Local Similarity: 48.76% Mismatches: 47
Query Match: 62.56% Indels: 140
DB: 4 Gaps: 5

US-09-827-854-18 (1-317) x US-08-949-155-51 (1-4267)
QY 1 MetLysValLeuThrPalaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2448 ATGGCTGTGACATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2489
QY 21 GluGlnAlaValGlnThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 37
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2490 GAGCGCGGGG-----CCGCGCGCGAGAGTGCACGTGTGTGGAGAGAGAGAGAGAG 2537
QY 38 TrpGlnSerGlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrp 57
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2538 TGGCAGGAGCAGCAGCCCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2597
QY 58 ValGlnThrLeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlu 77
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2598 GTGAGAGCCCTGCTGACCAAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2657
QY 78 Leu----- 78
Db 2658 CT-GACGTAAGTCCACCCGAGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCTGA 2716
QY 78 ----- 78
Db 2717 CCTCTCGGCAACCGTGTGTTGACACCTCAGAGCTCAGCCGCTCCGGGTTCTTCTG 2776
QY 78 ----- 78
Db 2777 TCCTGTGCGCAACTCTGGGGGCTGGGTCTGTCTGTTCTTTTCTTTCTCTCTCTT 2836
QY 78 ----- 78
Db 2837 TTGGGGGAAACATTTTCTTTTCTTTTCTTTCAATTGACTTCANGTCTGCTTCTTCC 2896
QY 78 ----- 78

```



```

Db 2837 TTGGGGGGAAGAAACTTTTCTTTCTTTCTTTCTTCAATTGACTTCATGCTCTTCTTCTTCC 2836
QY 78 ----- 78
Db 2897 ATCTTGAGCTCTCGCTGCTTTCGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2956
QY 78 ----- 78
Db 2957 AATCTCTGGCAGCTCTGGCCATTCGCCAGCTCAGAGACCTCTCTCTCTCCCTCAGCCGC 3016
QY 79 ----- ArgAlaLeuMetAspGluThrMetIysGluLeuIysAlaThr 92
Db 3017 CCGGCCCTCTCTGGCGCCAGGAGAGCTATGAGAGAGCATGAGAGAGCTGAGAGAGCTGAC 3076
QY 93 LysSerGluLeuGluGluGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSer 112
Db 3077 GCGCGAGAGCTGGAGAGCGGAGAGCTGGGCGCCGTGACCCAGAGAGAGAGCGAGCGGCTGTC 3136
QY 113 LysGluLeuGlnAlaAlaGlnAlaAlaArgLeuGlyAlaAspMetGluAspValCysIleArg 132
Db 3137 AAGGAGCTGCAGGGGGGCGAGGCCCGCGTGGGGCGGCAATGAGAGACCTGGCAACCGC 3196
QY 133 LeuValGlnThrArgArgGlyValAlaGlnAlaMetLeuGlyGlnSerThrGluLeuArg 152
Db 3197 TTGGTGCTCTACCCGACGAGAGTGCACACATTTTGGGCGACACACCGAGAGAGCTGGCG 3256
QY 153 ValArgLeuAlaSerHisLeuArgLysLeuCysLysArgLeuLeuArgAspAlaAsp 172
Db 3257 AGCGCCCTGGCTTCCACACTGCGCAGAGCTGCGCAACGGCTGCTCGGACACCGAGGAC 3316
QY 173 LeuGlnLysArgGluAlaValLysGlnAlaGlyAlaArgGluGlyAlaGlnArgLysLeu 192
Db 3317 CTGCGAAGAGCCCTGGCGCTGCTACAGAGGCGGGGCTCGCGAGGGCGCGCAGCGCAGCGTG 3376
QY 193 SerAlaIleArgGluArgLysGluGlyProLeuValGluGlnGlyArgValArgAlaIleThr 212
Db 3377 AGCGCCCTCGCGAGGCGCTCGGGCCCTGCTGTGAGACAGGCGCAATGGCGCCGCCAC 3436
QY 213 ValGlySerLeuAlaGlyGlnProLeuGlnGlnArgAlaGlnAlaThrPcylGluArgLeu 232
Db 3437 CTGAGTACAGAGCGCGGCGCAGCGCTGCGGACGCGCGGAGACCTGGGGCGCAGAACCTG 3496
QY 233 ArgAlaIleArgMetGluGluMetGlySerArgThrArgAspArgLeuAspGluValLysLeu 252
Db 3497 GCGGAGAGCGCTGGGAGGAGTGGGACCGGACCGCGGACCGGCTGTGATGAGATGGCTGAG 3556
QY 253 GlnAlaIleGluValArgAlaLysLeuGlnGlnAlaGlnAlaIleArgLeuGlnAla 272
Db 3557 CAGCTGAGAGAGAGGCGCCACCAAGTGGAGAGACAGGACGAGCAATGGCTGCGTCAAGCC 3616
QY 273 GlnAlaIleGlnAlaArgLeuLysSerThrPheGluProLeuValGluAspMetGlnArg 292
Db 3617 GAGGAGATTCCAGCCCTCTCAAAAGCTGGTTGAGACCTCTGGTGAAGACATACAGCGCC 3676
QY 293 GlnIlePheAlaGlyLeuValGlyLysValGlnAlaIleValGly---ThrSerAlaIlePro 311
Db 3677 CAGTGGCGCGGGGCTGGTGGAGAGATGACAGTGGCGGTGACATAACTCTCCACACTCT 3736
QY 312 ValProSerAspAsn 316
Db 3737 GCGGCCAGTGTATAT 3751

```

```

1 ADDRESSSEE: Banner & Mitcoff, Ltd.
2 STREET: 1 Financial Center
3 CITY: Boston
4 STATE: MA
5 COUNTRY: US
6 ZIP: 02111
7
8 COMPUTER READABLE FORM:
9 MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
10
11 COMPUTER: IBM PC compatible
12 OPERATING SYSTEM: PC-DOS/MS-DOS
13 SOFTWARE: Wordperfect 6.1
14
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER: US/08/726, 306A
17 FILING DATE: 02-Oct-1996
18
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: GB 95/20080.4
21 FILING DATE: 02-Oct-1995
22
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: US 60/009, 832
25 FILING DATE: 01-Jan-1996
26
27 ATTORNEY/AGENT INFORMATION:
28 NAME: Williams, Ph.D., Kathleen M.
29 REGISTRATION NUMBER: 34,380
30 REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
31
32 TELECOMMUNICATION INFORMATION:
33 TELEPHONE: (617) 345-9110
34
35 INFORMATION FOR SEQ ID NO: 28:
36
37 SEQUENCE CHARACTERISTICS:
38 LENGTH: 660 base pairs
39 TYPE: nucleic acid
40 STRANDEDNESS: double
41 TOPOLOGY: linear
42
43 MOLECULE TYPE: CDNA
44
45 US-08-726-306A-28
46
47
48
49 Alignment Scores:
50 Pred. No.: 1,18e-87 Length: 660
51 Score: 976.00 Matches: 196
52 Percent Similarity: 98.49% Conservative: 0
53 Best Local Similarity: 98.49% Mismatches: 3
54 Query Match: 61.42% Indels: 0
55 DB: 2 Gaps: 0
56
57 US-09-827-854-18 (1-317) x US-08-726-306A-28 (1-660)
58
59 Oy 1 MetLysValLeuTrpAlaAlaLeuIleuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
60 Db 62 ATGAGGATTCTGTGGGCTGCGCTGTGTCATTTCTCGCAGATGCCAGGCAAGGTG 122
61
62 Oy 21 GluGlnAlaValaGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
63 Db 122 GAGCAGCGGGTGGAGACAGACCCGGAGCCGAGCTGGCCGACGACGACCGAGTGGCAAGC 182
64
65 Oy 41 GlyGlnArgTrpGluLeuAlaLeuGlyLysArgPheTrpAspTrpLeuArgTrpValGlnThr 60
66 Db 182 GGCACGCGCTGGGAAGTGGCACTGGGTGCTTTGGGATTACCTGGCCGTGGGTGCAGACA 242
67
68 Oy 61 LeuSerGluGlnValGlnGluLeuLeuSerGlnValThrGlnGluLeuArgAla 80
69 Db 242 CTGTCTGAGCGAGGTGCAGAGAGACTGCTCAGCTCCCAAGTCAACCCAAAGAACTGAGGGCG 302
70
71 Oy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGluGlnLeu 100
72 Db 302 CTGATGAGCAGACCATGAAAGATTGAAGCCCTCAACAATCGGAACCTGAGAGAAACAATG 362
73
74 Oy 101 ThrProValaIaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaGlnAla 120
75 Db 362 ACCCGGTAGCGGAGAGAACCCGGCCAGCGCTGCTCAAGAGAGCTGCAGACGCGCAGGCC 422
76
77 Oy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
78 Db 422 CGGCTGGGCGCGGACATGAGAGAGATGTGGCGCGCTGTGTGCAATGCCGCGGAGGTG 482
79

```


Oy	109	AlArGLEuSerLySGluGlnAlAaladInIArgrLEuGLyAlAsPMetGIuASP	128
Dd	2	GCAAGCGCTGTCCAAAGACACTCAGACGGCGGCAGGCCGCTGGCGCGGACATGTGAAGAC	61
OY	129	VAlCySglArGrLeuValGlntYrArGlyGlLuVAIglInAlmeLleuGlYgInSeTThr	148
Dd	62	GTCGTCCG - CCGCTGTGTGCATTACC GCGCGGACAGGTGCAGGCCAATGCTCGGCCAAGACACC	120
OY	149	GIUGlUeuaArgValArGrLeuAlASeTHisLeuNrgLySteuCySyLSArGrLeuLeuArg	168
Dd	121	GAGGAGCTGC GG GTGGCTTCGCTCCCACTGTGGCAAAGCTGCTAAAGCGACTTCCTCCG	180
OY	169	ASpAlAAsPAPLeuGlnILNySarGrLeuAlAValTyrcInAlaGLyAlArGrGIuGLyAlA	188
Dd	181	GATGCCGATGACCTGCAGAACTCCCTGGCACTGTACCAAGCGCGGGCCCGCGAAGGGCC	240
OY	189	GIUAArgGIuLeu	192
Dd	241	GAGCGCGGGCTTC	252

MOLECULE TYPE: CDNA		US-09-431-613-24	
Alignment Scores:			
Pred. No.:	3.87e-28	Length:	252
Score:	366.00	Matches:	81
Percent Similarity:	96.43%	Conservative:	0
Best local Similarity:	96.43%	Mismatches:	3
Query Match:	23.03%	Indels:	1
DB:	4	Gaps:	0
US-09-827-854-18 (1-317) x US-09-431-613-24 (1-252)			
QY	109	AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGluAlaAspMetGluAsp	128
Db	2	GCAGCGGTGTCCAAGAGCATGACAGCGGAGCGAGCCGCGGTGGCGCGGACATGAGAGAC	61
QY	129	ValCysGlyArgLeuValAlaGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr	148
Db	62	GTGTGCGCG-CGCTGCTGTGCTACCTACCGCGGAGGTGAGGCGCATGCTGGCGAGAGCACC	120
QY	149	GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuCysLysArgLeuLeuArg	168
Db	121	GAGGAGGTGCGGCGGTGGCGCTGCTCCACCTGCGCAAGCTGGCTAGCGGCTCTCGC	180
QY	169	AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla	188
Db	181	GATGCCGATGACCTGCAGAAATCCCTGTCAGTGTACAGGCGGCGGCGCGAGGGCGCC	240
QY	189	GluArgGlyLeu	192
Db	241	GAGCGCGGCTCTC	252
RESULT 11			
: Sequence 24, Application US/09504245			
: Patent No. 6221605			
GENERAL INFORMATION:			
APPLICANT: K Ster, Hubert			
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry			
NUMBER OF SEQUENCES: 33			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Heller Ehrman White & McAlliff LLP			
STREET: 4250 Executive Square, 7th Floor			
CITY: La Jolla			
STATE: CA			
COUNTRY: USA			
ZIP: 92037-9103			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Diskette			
COMPUTER: IBM Compatible			
OPERATING SYSTEM: DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/09/504,245			
FILING DATE:			
CLASSIFICATION:			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: 08/617,256			
FILING DATE: 18-MAR-1996			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: 08/406,199			
FILING DATE: 03-MAR-1995			
ATTORNEY/AGENT INFORMATION:			
NAME: Seidman, Stephanie L			
REGISTRATION NUMBER: 33,779			
REFERENCE/DOCKET NUMBER: 24736-2002J			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: 858-450-8400			
TELEFAX: 858-587-5360			
INFORMATION FOR SEQ ID NO: 24:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 252 base pairs			

```

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-504-245-24

Alignment Scores:
Pred. No.:      3.87e-28      Length:      252
Score:          366.00        Matches:     81
Percent Similarity: 96.43%    Conservative: 0
Best Local Similarity: 96.43% Mismatches:     3
Query Match:     23.03%       Indels:      1
DB:              4           Gaps:        0

US-09-827-854-18 (1-317) x US-09-504-245-24 (1-252)

Oy      109 AlArGrLeuSerLySGluLeuGlnAlAAGlInAlArGrLeuGlnLyAlAsPmetGlUAsP 128
      |||||||
Db      2 GCACGGCGTGTCCAAGAGACTTCAGAGCCGCCAGGCCCGCTGGGCCGGGCATGGAAGAC 61

Oy      129 ValCgSLyArGrLeuValGlnTyArGlyGLyuValGlnAlnAmelLeuGlnInsTrHr 148
      |||||||
Db      62 GTGGGCGC -CGCTGGTGCACTACACGGCGGAGGTGCAGGCCCATMGCTCGGCCAGAGACC 120

Oy      149 GluGUleuArGVAlArGrLeuAlAserHisLeuArGlyIsLeucYslySArGrLeuLeuArG 168
      |||||||
Db      121 GAGAGAGCTGCGGGGCGCCCTCCCTCCACCTGGCAAGCTGGCTAAAGCGGCTCTCCCG 180

Oy      169 AsPlAAsPaSPleuGlnLyASrArgLeuAlValTyGlnAlnAglyAlArGrGluGIyAlA 188
      |||||||
Db      181 GATCCCATGACTCTGCACAGACTCCCTTCGCACTGTACCAAGCCCGGGCCCGCAGAAGCGCC 240

Oy      189 GlUArGrGlyLeu 192
      |||||||
Db      241 GAGCGGCGCTC 252

RESULT 12
US-09-287-682-24
Sequence 24, Application US/09287682
Patent No. 6235478
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
APPLICANT: Braun, Andreas
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McAniff
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002E
TELECOMMUNICATION INFORMATION:

```


TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-682-24

Alignment Scores:
Pred. No.: 3.87e-28 Length: 252
Score: 366.00 Matches: 81
Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 23.03% Indels: 1
Gaps: 0

US-09-827-854-18 (1-317) x US-09-287-682-24 (1-252)

QY 109 AAlarqleuserlysgluleuGlnAlaAlaGlnAlaargleuGlyAlaaspmetGluasp 128
Db 2 GCACGGCTGTCCAGAGAGCTGCAGCGCGCGCGCGCGCGCGCATGAGAGAC 61

QY 129 ValcysglYarGleuValGlnTyrArgGlyGluValGlnAlaMetleuGlyGlnSerThr 148
Db 62 GTGTGCGC-CGCTGTGTGCTACCGCGCGCGAGGTGCGAGGCGGCGGCGGCGAGAC 120

QY 149 GluGluLeuArGValArGleuAlaSerHisLeuArgLysLeuCysLysArGleuLeuArg 168
Db 121 GAGAGCTGGGGGTGGCGCTGCGCTCCACCTGCGCAAGCTGCGTAAGCGGCTCTCCGC 180

QY 169 AspaAlaaspLeuGlnLysArGleuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
Db 181 GATCCCATGACCTGCAGAGAGTCCCTGCGAGTGTACAGGCGCGGCGCGCGAGGCGCC 240

QY 189 GluArgGlyLeu 192
Db 241 GAGCGGCGCCTC 252

RESULT 13
US-09-287-679-24
Sequence 24, Application US/09287679
Patent No. 6258538
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McAuliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,679
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-679-24

Alignment Scores:
Pred. No.: 3.87e-28 Length: 252
Score: 366.00 Matches: 81
Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 23.03% Indels: 1
Gaps: 0

US-09-827-854-18 (1-317) x US-09-287-679-24 (1-252)

QY 109 AAlarqleuserlysgluleuGlnAlaAlaGlnAlaargleuGlyAlaaspmetGluasp 128
Db 2 GCACGGCTGTCCAGAGAGCTGCAGCGCGCGCGCGCGCGCATGAGAGAC 61

QY 129 ValcysglYarGleuValGlnTyrArgGlyGluValGlnAlaMetleuGlyGlnSerThr 148
Db 62 GTGTGCGC-CGCTGTGTGCTACCGCGCGCGAGGTGCGAGGCGGCGGCGGCGAGAC 120

QY 149 GluGluLeuArGValArGleuAlaSerHisLeuArgLysLeuCysLysArGleuLeuArg 168
Db 121 GAGAGCTGGGGGTGGCGCTGCGCTCCACCTGCGCAAGCTGCGTAAGCGGCTCTCCGC 180

QY 169 AspaAlaaspLeuGlnLysArGleuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
Db 181 GATCCCATGACCTGCAGAGAGTCCCTGCGAGTGTACAGGCGCGGCGCGCGAGGCGCC 240

QY 189 GluArgGlyLeu 192
Db 241 GAGCGGCGCCTC 252

RESULT 14
US-09-397-766-24
Sequence 24, Application US/09397766
Patent No. 6268144
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McAuliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/397,766
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-20021
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-397-766-24

Alignment Scores:
Pred. No.: 3,87e-28
Score: 366.00
Percent Similarity: 96.43%
Best Local Similarity: 96.43%
Query Match: 23.03%
DB: 4
Length: 252
Matches: 81
Conservative: 0
Mismatches: 3
Indels: 1
Gaps: 0

US-09-827-854-18 (1-317) x US-09-397-766-24 (1-252)

OY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGGGCGGCGGCGGCTGGGCGCGGACATGGAGGAC 61
OY 129 ValCysGlyArgLeuValGlnTyrArgGlyValAlaGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGGCC-CGCTGTGTGAGTACCGCGGCGAGGTGCAGGCATGCTCGGCCAGAGACACC 120
OY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuGlnCysLysArgLeuLeuArg 168
DB 121 GAGGAGCTCGGGGTGGCTCTCCACCTGCGCAAGCTGCTGAAGCGGCTCTCCCGC 180
OY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAAATCCCTGCGCATGTACAGCGGGGCGCGGAGGGCGCC 240
OY 189 GluArgGlyLeu 192
DB 241 GAGCGGGGCGCTC 252

RESULT 15
US-09-287-681-24

; Sequence 24, Application US/09287681
; Patent No. 6277573
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037-9103

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,681
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-287-681-24

Alignment Scores:
Pred. No.: 3,87e-28
Score: 366.00
Percent Similarity: 96.43%
Best Local Similarity: 96.43%
Query Match: 23.03%
DB: 4
Length: 252
Matches: 81
Conservative: 0
Mismatches: 3
Indels: 1
Gaps: 0

US-09-827-854-18 (1-317) x US-09-287-681-24 (1-252)

OY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGGGCGGCGGCGGCTGGGCGCGGACATGGAGGAC 61
OY 129 ValCysGlyArgLeuValGlnTyrArgGlyValAlaGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGGCC-CGCTGTGTGAGTACCGCGGCGAGGTGCAGGCATGCTCGGCCAGAGACACC 120
OY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuGlnCysLysArgLeuLeuArg 168
DB 121 GAGGAGCTCGGGGTGGCTCTCCACCTGCGCAAGCTGCTGAAGCGGCTCTCCCGC 180
OY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAAATCCCTGCGCATGTACAGCGGGGCGCGGAGGGCGCC 240
OY 189 GluArgGlyLeu 192
DB 241 GAGCGGGGCGCTC 252

Search completed: March 14, 2003, 20:18:51
Job time : 39.7343 secs

Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

QY	1	MetylsValLeuTfPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal	20
Db	61	ATGAAGGTTCTGTGGCGCTTCTCTGTCAATCTCTGGCAGATCCAGGCCAAGTC	120
QY	21	GlUcInAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluProGlnSer	40
Db	121	GAGCAAGCGGTGGAGACAGACCAGGACCCGAGCTGGCCAGCAGACCGAGTGGCAGAGC	180
QY	41	GlyGlnArgTfPgluLeuAlaLeuGlyArgPheTrpAspIyrLeuAArgTfPvalGlnThr	60
Db	181	GGCCAGCGCTGGAACTGGCACTGGGTCTCTTGGGATTAACCTGGCCCTGGGTGCAGACA	240
QY	61	LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80
Db	241	CTGTCTGAGCAGGTGGCAGAGAGACTCTCTCAGCTCCAGGTCCAGCAGCAGAACTGAGGGCG	300
QY	81	LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu	100
Db	301	CTGATGGACGACACATGAAAGATTGAAGCCCTACAAATCGCAACTGGAGGAAACAACCTG	360
QY	101	ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	361	ACCCGGGGGGGGAGAGACCGGGCACGGGTCTCCAAAGAGCTGGCAGGGGGCGCAGGCC	420
QY	121	ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal	140
Db	421	CGGCTGGGGCGGACATGAGAGAGACTGTGGGGCCCGCTGGTGCACTACCGCGGGCAGGTC	480
QY	141	GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	481	CAGGCCATGCTCGGGCCAGAGCACCGAGAGACTGGGGGTGGCTCGCTGCCCTCCACCTGGCC	540
QY	161	LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	541	AAGCTGTGTAAAGCGGCTCTCTCGGGATGCGGATGACTCTGCAGAAAGCCCTGGCGAGTGTAC	600
QY	181	GlnAlaGlyAlaArgGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200
Db	601	CAGCGCGGGGCCCGCAGAGGGGGCGGAGCGGGCTCTCAGCCCAATCCGCGAGCGCTGGGG	660
QY	201	ProLeuValGlnGlnGlnArgValArgAlaAlaThrValIleLysLeuAlaGlnPro	220
Db	661	CCCCGTGGGAAACAGGGCCGGGTGCGGGGCCCACTGTGGGCTCTCCGTGGCGGCCACCGC	720
QY	221	LeuGlnGluArgAlaGlnAlaTfPgluArgLeuArgAlaArgMetGlnGluMetGly	240
Db	721	CTACAGAGAGCGGGCCAGAGCTTGGGGGCGAGCGCTGGCGCGCGATGGAGAGATTGGGC	780
QY	241	SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys	260
Db	781	AGCGGAGCCCGCAGACCGGCTGGAGAGAGTGAAGAGCAGAGTGGCGGAGGTGGCGGCCAAG	840
QY	261	LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys	280
Db	841	CTTGAGAGAGCGGCCCGCAGATTCAGGCTCTCAGGCGCAGGCTCTTCCAGGCGCCGCTCAG	900
QY	281	SerTrpPheGluProLeuValGluAspMetGluArgGlnTfPgluGlyLeuValGlyLys	300
Db	901	AGCTGGTTCAGAGCCCTGGTGGTGAAGACATGACAGCCAGTGGGGCGGGGCTGGTGGAGAG	960
QY	301	ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis	317
Db	961	GTGCAGAGCTGGCGTGGGCGACAGAGCGCCCGCTGTGTGCCAGCACAATATAC	1011

```

: Patent No. US20020177551A1
: GENERAL INFORMATION:
: APPLICANT: TERMAN, David S
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
: FILE REFERENCE: 8707759
: CURRENT APPLICATION NUMBER: US/09/870,759
: CURRENT FILING DATE: 2002-01-14
: PRIOR APPLICATION NUMBER: US 60/208,128
: PRIOR FILING DATE: 2000-05-30
: NUMBER OF SEQ ID NOS: 166
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 129
: LENGTH: 1156
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (61)..(1014)
: OTHER INFORMATION:
: US-09-870-759-129

```

Alignment Scores:			
Pred. No.:	6.07e-137	Length:	1156
Score:	1577.00	Matches:	316
Percent Similarity:	99.68%	Conservative:	0
Best Local Similarity:	99.68%	Mismatches:	1
Query Match:	99.24%	Indels:	0
DB:	9	Gaps:	0
US-09-827-854-18 (1-317) x US-09-870-759-129 (1-1156)			
QY	1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20	
DB	61 ATGAAGTTCGTGGGGCTGCTGCTGGTACATTCTCGAGAGATCCAGGCCAAGTGG	120	
QY	21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnInnThrGluTrpGlnSer	40	
DB	121 GAGCAAGCGGTGGAGACAGACCCGAGCCGACGTGCGCCAGCAGACCGAGTGGCAGAGC	180	
QY	41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr	60	
DB	181 GGGCAGCGCCCTGGAAACTGGCACTGGGTGCGCTTTGGGATTACCTGGCCTGGGTGCACACA	240	
QY	61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80	
DB	241 CTGTCTGAGCAGGTGCAGAGAGAGACTGCTCAGTCCCGAGGTACCCAGAACTAGAGGGCG	300	
QY	81 LeuMetAspGluThrMetLeuGluLeuLysAlaTrpLysSerGluLeuGluGluGlnLeu	100	
DB	301 CTGATGGACGAGACCATGTAAAGATTGAAGGCTTCAAAATCGGAACGTGGAGAAACACTG	360	
QY	101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120	
DB	361 ACCCGGCGGGGAGAGAGACCGGGGGCACGGCTGTCCAAAGACGTGCAGCGGGCCAGAGCC	420	
QY	121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal	140	
DB	421 CGGCTGGGGCGCGACATGTAGAGACTGTGGCGGGCGCGTGGGCACTACCGCGGGAGAGTG	480	
QY	141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg	160	
DB	481 CAGGCAATGCTCGGCGCCAGAGCACCGAGAGAGCTCGGGGTGGCCCTCGCTCCACCTGGCG	540	
QY	161 LysLeuSerLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180	
DB	541 AAGTGCCTAAGCGCGCTCTCTCCGATGCCGATACCTGCAGAAAGCCCTGGCAGTGTAC	600	
QY	181 GlnAlaGlyAlaAspGlyGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200	
DB	601 CAGGCGGGGGCGCCGAGAGCGCGCGAGCGGCTTACGCGCAATCCGCGAGCGGCTTGGG	660	
QY	201 ProLeuValGluGlnGlyArgValAlaArgAlaAlaThrValGlySerLeuAlaGlyInPro	220	

Db	661	CCCTCGTGGAAACAAGGCGCGCTCGGGCGCCACCTGTGGGCTCCCTGGCGGGCAACCC	720
QY	221	LeuGlnGluIuaTgAlaGlnAlaITrpGlyGluIuaTgAlaTgMetGluGluMetGly	240
Db	721	CTACAGAGAGCGGGCCACAGCGCTGGGGGCGAGCGCTGGCGCGGAGAGAGAGATGGGC	780
QY	241	SerATgTrrTgTgAsPArgLeuAsPglValIysGluGlnAlaGluValATgATgAlaTys	260
Db	781	AGCGGACCCCGCGACCGGCTGTGACAGAGTGAAGAGCAGGTGGCGAGATGCGCCGACAG	840
QY	261	LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaTgLeuLys	280
Db	841	CTGGAGGAGCGAGGGCCACACATCGGCTGAGAGCCAGCGCGCTTCCAGGCGCGCTCAAG	900
QY	281	SerTrrPheGluPArgProLeuValGlnAsPmetGlnIArgGlnTrrPAlaGlyLeuValGluLys	300
Db	901	AGCTGGTTGCGAGCCCGTGTGTGGAACATGACAGGCCCAATGTGGCGGCGTGTGTGAGAAC	960
QY	301	ValGlnAlaIaIaValGlyThrSerAlaIaProValProSerAspAsnHis	317
Db	961	GTTCAGAGGTGCGGTGGGACACAGCGCGCGCCCTGTGTGCCACAGCAATATCAC	1011

RESULT 3

```

US-09-802-640-17
: Sequence 17, Application US/09802640
: Publication NO. US20030036057A1
:
: GENERAL INFORMATION:
:
: APPLICANT: Braun, Andreas
: APPLICANT: Bonsal Aruna
: APPLICANT: Kleyan Patrick
: TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: CAROTIOVASCULAR DISEASE AND THEIR USE
: FILE REFERENCE: 24736-2048
: CURRENT APPLICATION NUMBER: US/09/802,640
: CURRENT FILING DATE: 2001-03-09
: NUMBER OF SEQ ID NOS: 122
: SOFTWARE: FastSeq for Windows Version 4.0
:
: SEQ ID NO 17
:
: LENGTH: 1156
:
: TYPE: DNA
:
: ORGANISM: Homo sapien
:
: FEATURE:
:
: NAME/KEY: CDS
:
: LOCATION: (61)...(1014)
:
: OTHER INFORMATION: Nucleotide Sequence encoding apolipoprotein E
:
: US-09-802-640-17

```

Alignment Scores:

Pred. No.:	6.07e-137	length:	1156
Score:	1577.00	Matches:	316
Percent Similarity:	99.688	Conservative:	0
Best Local Similarity:	99.688	Mismatches:	1
Query Match:	99.248	Indels:	0
DB:	9	Gaps:	0

US-09-827-854-18 (1-317) x US-09-802-640-17 (1-1156)

Qy	1	MelystValLeuThrPAlaAlaLeuLeuValThrPheLeuAlaGlyGsgInalAlaVal	20
Db	61	ATGAAGGTCCTGTGGGCTGGCTCTGTCACATCTCTGGCAGAGATCCAGGGCAAGGTC	120
Qy	21	GlulnAlaValAlaGluThrGluProGluProGluLeuArgInaInThrGluTrpGlnSer	40
Db	121	GAGCAAGCGGTGGAGACACAGACCCGAGCTGGCCACACAGACCCGAGTGGCAAGC	180
Qy	41	GlyIleArgTrpGluLeuAlaLeuGluAlaRphenTrpAspTrpLeuArgTrpAlaGlnThr	60
Db	181	GGCCAGCGCTGGGAACGTGGACCTGGGTGGCTTTGGGATTACCTGGCTGGGTGGACACA	240
Qy	61	LeuSerGluGlnValGlnGluGluLeuLeuSerGlnValThrGlnGluLeuArgAla	80
Db	241	CTGTCTGGACGAGTGGAGAGAGACTCTCAAGCTCCAGAGTACACCCAGGAATGAGGGCG	300

QY	81	LeuNetAspArgLupTrmMetLysGluLeuLysAlaIarTrpSerGluLeuGluGluLeu	100
Dp	301	CTGATGGACGAGACCATTAAAGAAATTGAAGGCTTCAAAATCGGAATCGAGGAACAAC	360
QY	101	ThrProValAlaGluGluGluThrArgAlaArgLeuSerTrpGluLeuGlnAlaAlaGlnAla	120
Dp	361	ACCCCGGTGGCGGAGAGACACCGGGCACGGCTGTTCAGAGAGCTGCAGGGCGCCAGGCC	420
QY	121	ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal	140
Dp	421	CGGCTGGGGCGGAGCATGGAGAGAGCTGTGGCGCGCTCGTGTGCATCCGGGGGAGAGTG	480
QY	141	GlnAlaMetLeuGlyGlnSerThrGluGluLeuAArgValArgLeuAlaSerHisLeuArg	160
Dp	481	CAGGCATGCTCTGGCCACAGACACCGAGAGAGCTGGGGGTCCTCGGCTCCACCTCGCG	540
QY	161	LysLeuCysLysAspGluLeuAArgAspAlaAspAspLeuGlnLysAspGluLeuAlaValTyr	180
Dp	541	AAGGCGCTAAGCGGCTCTCCGGGATGCCCATATCCTGCAGGAAGCCCTGGCGAGTGTAC	600
QY	181	GlnAlaGlyAlaAspGluGluGlyAlaGluAArgGlyLeuSerAlaIleArgGlyAArgLeuGly	200
Dp	601	CAGCGCGGGGCCCGCGAGAGGGCGCCGAGAGCGGCTCTCAGGCCATCCGCGAGCGCTGGGG	660
QY	201	ProLeuValGluGluGlnGlyArgValAArgAlaAlaThrValGlySerLeuAlaGluGlnPro	220
Dp	661	CCCTGGGTGAACAAGGGCGCGTGCGGGGCCGCACCTGTGGGCTCCGTGGCGGGCACCGG	720
QY	221	LeuGluGluAArgAlaGlnAlaIleTrpGlyGluAArgLeuAArgAlaArgMetGluGluMetGly	240
Dp	721	CTACAGGAGCGCGGCCAGAGGCTGTGGGGCGAGCGCTGGCGCGCGAGATGAGATCGGC	780
QY	241	SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAArgAlaLys	260
Dp	781	AGCCGGACCCCGACCGCTGTGACGAGGTGAAGAAGCGAGGTGGGAGGTGCGCGCCAG	840
QY	261	LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys	280
Dp	841	CTGGAGGAGCAGGCCCGACCAATATACGCTGCAGGGCCGAGGCTTCCAGGCCCGGCTCAAG	900
QY	281	SerTrpPheGluProLeuValGluAspMetGlnAArgGlnTrpAlaGlyLeuValGluLys	300
Dp	901	AGCTGGTTCGAGCCCTGTGGTGAAGACATGCAGCGCCAGTGGCGGGGCTGTGTGAAGAG	960
QY	301	ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis	317
Dp	961	GTTCAGAGGCTGCCGTGGGCACCGAGCGCCCTGTGTGCCACAGCGCAATATCAC	1011

RESULT 4

```

US-09-827-854 8
: Sequence 8, Application US/09827854
: Patent No. US20020123093A1
:
: GENERAL INFORMATION:
: APPLICANT: Zannis, Vassilis
: TITLE OF INVENTION: Compounds and methods for lowering
: TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia
: FILE REFERENCE: 07180/004003
: CURRENT APPLICATION NUMBER: US/09/827,854
: CURRENT FILING DATE: 2001-04-05
: PRIOR APPLICATION NUMBER: US 09/679,088
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 09/544,386
: PRIOR FILING DATE: 2000-04-06
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 8
: LENGTH: 1156
:
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-827-854-8

```


Alignment Scores:

Pred. No.:	6,07e-137	Length:	1156
Score:	1577.00	Matches:	316
Percent Similarity:	99.68%	Conservative:	0
Best Local Similarity:	99.68%	Mismatches:	1
Query Match:	99.24%	Indels:	0
DB:	10	Gaps:	0

US-09-827-854-18 (1-317) x US-09-827-854-8 (1-1156)

```

OY 1 MetLysValLeuTrpAlaIalaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||||||
DB 61 ATGAAGGTTCTGTGGCTCGTTCGCTGACATTCTTGGCAGAGATGCCAGGCCAAGGTG 120
OY 21 GluGlnAlaValaGluThrLupProGluLeuArgGlnGlnThrGlnTrpGlnSer 40
    |||||||
DB 121 GAGCAAGCGGTGGAGACAGAGCCGAGCTGCGCAGACAGCCAGTGGCGAGAGC 180
OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLysLeuArgTrpValGlnThr 60
    |||||||
DB 181 GGCCAGCGCTGGGAACCTGGCTGGCTGCTTTGGGATTACCTGGCTGGGTGGCAGACA 240
OY 61 LeuSerGlnGlnValaGlnGluLeuLeuSerSerGlnValaThrGlnGluLeuArgAla 80
    |||||||
DB 241 CTGTCTGAGCAGAGTGCAGAGAGAGCTGCTCAGCTCCAGAGTACCAGGAACTGAGGGCG 300
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGlnLeuGluGlnLeu 100
    |||||||
DB 301 CTGATGACAGACACATGAAAGAGTTGAAGCGCTTACAAATCGGAACCTGAGGAACTG 360
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
    |||||||
DB 361 ACCCGGCTGGGAGAGAGAGCGGGCGACGCTGTCGAAGAGCTCGAGCGCGCGAGCGCC 420
OY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
    |||||||
DB 421 CGGCTGGCGCGGACATGAGAGAGCTGTGCGGCCCTGCTGCTCAGTACCGCGCGAGGTG 480
OY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValaArgLeuAlaSerHisLeuArg 160
    |||||||
DB 481 CAGGCATGCTCGGCGCAGAGACCGGAGAGCTGCGGGTGCCTGCGCTGCCACCTGGCGCC 540
OY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValaLys 180
    |||||||
DB 541 AAGCTGCGTAAACGCGCTCTCCGCGATGCGGATGACCTGCAGAACGCGCTGGCAGGTAC 600
OY 181 GlnAlaGlyAlaArgGluGluValaGluArgGlyLeuSerAlaIleArgGlyArgLeuGly 200
    |||||||
DB 601 CAGGCGCGGGCGCGGAGGGCGCCGAGCGCGCTCAGCGCATCCGCGAGCGCTTGGGG 660
OY 201 ProLeuValaGluGlnGlyArgValaArgAlaAlaThrValaGlySerLeuAlaGlyGlnPro 220
    |||||||
DB 661 CCCCTGGTGAAGAGAGCGCGCTGGCGGGCGCCACTGTGGGTCTCCCTGGCGCGGACGCG 720
OY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
    |||||||
DB 721 CTACAGAGAGCGGGCGGAGCGGCGGCTGCGCGCGCGGATGATGAGAGATGGGC 780
OY 241 SerArgThrArgAspArgLeuAspGluValaLysGlnGluValaGlnValaArgAlaLys 260
    |||||||
DB 781 AGCCGAGCGCGCGAGCGCTGGAGAGAGGTGAAGAGACAGGTGGCGAGGTGGCGCCAGAG 840
OY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
    |||||||
DB 841 CTGGAGAGAGAGCGCGCAGAGATACGCTGCGAGCGCGCTTCAGAGCGCGCTTCAAG 900
OY 281 SerTrpPheGluTrpLeuValaGluAspMetGlnArgGlnTrpAlaGlyLeuValaGluLys 300
    |||||||
DB 901 AGCTGCTTGAAGCGCGCTGCTGGAAGACATGACAGCGCCAGTGGCGCGGCTGGTGAAGAG 960
OY 301 ValGlnAlaAlaValaGlyThrSerAlaAlaProValProSerAspAsnHis 317
    |||||||
DB 961 GTGAGAGGCTGCGGTGGGACAGCGCGCGCTGTGGCCAGGCAACAAATCAC 1011

```

RESULT 5

US-10-044-090-454

Sequence 454, Application US/10044090

Patent No. US20020137081A1

GENERAL INFORMATION:

APPLICANT: Olga Bandman

TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION

FILE REFERENCE: PA-0028 US

CURRENT APPLICATION NUMBER: US/10/044,090

CURRENT FILING DATE: 2002-01-09

NUMBER OF SEQ ID NOS: 850

SOFTWARE: PERL Program

SEQ ID NO 454

LENGTH: 1291

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

OTHER INFORMATION: Incyte ID No. US20020137081A1 2514988CB1

LOCATION: 46

OTHER INFORMATION: a, t, c, g, or other

US-10-044-090-454

Alignment Scores:

Pred. No.:	6.94e-137	Length:	1291
Score:	1577.00	Matches:	316
Percent Similarity:	99.68%	Conservative:	0
Best Local Similarity:	99.68%	Mismatches:	1
Query Match:	99.24%	Indels:	0
DB:	12	Gaps:	0

US-09-827-854-18 (1-317) x US-10-044-090-454 (1-1291)

```

OY 1 MetLysValLeuTrpAlaIalaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||||||
DB 186 ATGAAGGTTCTGTGGCTCGTTCGCTGACATTCTTGGCAGAGATGCCAGGCCAAGGTG 245
OY 21 GluGlnAlaValaGluThrLupProGluLeuArgGlnGlnThrGlnTrpGlnSer 40
    |||||||
DB 246 GAGCAAGCGGTGGAGACAGAGCGGAGCGCGAGCTGCGCGCAGACAGCAGAGTGGCAGAGC 305
OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLysLeuArgTrpValGlnThr 60
    |||||||
DB 306 GGCCAGCGCTGGGAACCTGGCACTGGCTGCTTGGGATTACCTGCGCTGGGTGCAGACA 365
OY 61 LeuSerGlnGlnValaGlnGluGluLeuLeuSerSerGlnValaThrGlnGluLeuArgAla 80
    |||||||
DB 366 CTGTCTGAGCAGAGTGCAGAGAGAGCTGCTCAGCTCCAGGTACCCAGGAACTGAGGGCG 425
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnGlnLeu 100
    |||||||
DB 426 CTGATGAGAGAGACCATGAGAGAGTTGAAGGCTTCAAAATCGGAACCTGAGAGCAACTG 485
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
    |||||||
DB 486 ACCCGGCTGGGAGAGAGACCGGCGCAGCGCTGTCCAGAGAGCTGCAGCGCGCGAGGCC 545
OY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValaGlnTrpArgGlyGluVal 140
    |||||||
DB 546 CGGCTGGCGCGGACATGAGAGAGCTGTGGCGCGCGCTGTGTCACTACCGCGCGAGGTG 605
OY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValaArgLeuAlaSerHisLeuArg 160
    |||||||
DB 606 CAGGCATGCTCGGCGCAAGACAGCAGAGAGCTGCGGGTGGCGCTTCCCTCCACTGCGGC 665
OY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValaLys 180
    |||||||
DB 666 AAGCTGCGTAAAGCGGCTCTCCGCGATGCGGATGACCTGCAGAAAGCGCTGGCAGTGTAC 725
OY 181 GlnAlaGlyAlaArgGluGluValaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
    |||||||

```


US-09-827-854-9

Alignment Scores:

Pred. No.:	3,32e-136	Length:	1156
Score:	1569.00	Matches:	315
Percent Similarity:	99.37%	Conservative:	0
Best Local Similarity:	99.37%	Mismatches:	2
Query Match:	98.74%	Indels:	0
DB:	10	Gaps:	0

US-09-827-854-18 (1-317) x US-09-827-854-9 (1-1156)

```

QY 1 MetlyValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||||
DB 61 ATGAAGTTCTGTGGCTGGCTGTCTGTGCATTCCTGGCAGAGATCCAGGCAAGGTG 120

QY 21 GluGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGlnThrGlnSer 40
    |||||
DB 121 GAGCAACGGGTGGAGAGAGCCGAGCCGAGCTGCGCCAGCAGACCGAGTGGCAGAGC 180

QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
    |||||
DB 181 GGGCAGCGCTGGAGACCTGGCTGGCTTTGGGATTAACCTGGCTGGCTGGCTGACACA 240

QY 61 LeuSerGluGlnValGlnGluLeuLeuSerGlnValThrGlnGluLeuArgAla 80
    |||||
DB 241 CTGCTGAGCAGGTGCAGAGAGACTGCTCAGCTCCAGGTCACCGAAGTGAAGGCG 300

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
    |||||
DB 301 CTGATGAGCAGACCAAGAAAGAGTGAAGCCCTACAAATCGGAACGAGGAACAACGTG 360

QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
    |||||
DB 361 ACCCGGTGGCGAGAGAGAGCGGGGACGGCTGTCCAAAGAGCTGCAGCGCGCCAGGCC 420

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
    |||||
DB 421 CGGCTGGCGCCGACATGAGAGAGCTGTGCGCGCTGTGGTAGTACCCGCGGAGAGTG 480

QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
    |||||
DB 481 CAGCCATGCTCGGCCAGACACCGAGAGACTGGGGTGGCGCTCCGCTCCACACTCGC 540

QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
    |||||
DB 541 AAGCTGGCTAAGGGGCTCCCTCCGGATGCCGATACCTGCAGAAAGTCCCTGGCAGCTTAC 600

QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
    |||||
DB 601 CAGCCCGGGGCGCGAGGGGCGCGAGCGCGGCTCAGCCCATCCCGCAGCGCCCTGGG 660

QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValCysLysLeuAlaGlyGlnPro 220
    |||||
DB 661 CCCCTGGTGAACAGGGGCGCGTGGCGGCCCACTGTGGCTCCCTGGCGCGGACACCG 720

QY 221 LeuGlnLysArgAlaGlnAlaTyrGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
    |||||
DB 721 CTACAGAGAGGGGCGCGGCTGGGGGAGCGGCTGGCGCGCGGAGTGAAGATGGGC 780

QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
    |||||
DB 781 AGCGGAGCCCGCAGCCCTCGAGCAGAGTGAAGAGAGTGGCGGAGTGGCGGCGCAAG 840

QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
    |||||
DB 841 CTGGAGAGAGAGCGCCAGCAGATACGCTCGACAGCCGAGGCTTCACAGGCCGCTCAAG 900

QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTyrAlaGlyLeuValGluLys 300
    |||||
DB 901 AGCTGGTTCAGAGCCCTCGGTGGAAGACATGCAGCGCCAGTGGCGGCGTGGTGAAG 960

QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317

```

```

DB 961 GTGCAGGCTGCGGTGGACACAGCGCCCTGTGTGCCAGCAGCAATTCAC 1011
    |||||

```

RESULT 8

US-09-954-456-760

Sequence 760, Application US/09954456

Patent No. US20020115057A1

GENERAL INFORMATION:

APPLICANT: Young, Paul

TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C

FILE REFERENCE: 689290-76

CURRENT APPLICATION NUMBER: US/09/954,456

CURRENT FILING DATE: 2001-09-18

PRIOR APPLICATION NUMBER: US/60/233,617

PRIOR FILING DATE: 2000-09-18

PRIOR APPLICATION NUMBER: US/60/234,052

PRIOR FILING DATE: 2000-09-20

PRIOR APPLICATION NUMBER: US/60/234,923

PRIOR FILING DATE: 2000-09-25

PRIOR APPLICATION NUMBER: US/60/235,134

PRIOR FILING DATE: 2000-09-25

PRIOR APPLICATION NUMBER: US/60/235,637

PRIOR FILING DATE: 2000-09-26

PRIOR APPLICATION NUMBER: US/60/235,638

PRIOR FILING DATE: 2000-09-26

PRIOR APPLICATION NUMBER: US/60/235,711

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: US/60/235,720

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: US/60/235,840

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: US/60/235,863

PRIOR FILING DATE: 2000-09-27

NUMBER OF SEQ ID NOS: 2276

SOFTWARE: PatentIn version 3.0

SEQ ID NO 760

LENGTH: 1157

TYPE: DNA

ORGANISM: Homo sapiens

US-09-954-456-760

US-09-827-854-18 (1-317) x US-09-954-456-760 (1-1157)

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-827-854-18 (1-317) x US-09-954-456-760 (1-1157)

QY 1 MetlyValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20

DB 62 ATGAAGTTCTGTGGCTGGCTGTCTGTGCATTCCTGGCAGAGATCCAGGCAAGGTG 121

QY 21 GluGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGlnThrGlnSer 40

DB 122 GAGCAACGGGTGGAGAGAGCCGAGCCGAGCTGGCGCAGACGAGACCGAGTGGCAGAGC 181

QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60

DB 182 GGGCAGCGCTGGAGAGAGCTGGCTGGCTTTGGGATTAACCTGGCTGGCTGGCTGACACA 241

QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerGlnValThrGlnGluLeuArgAla 80

DB 242 CTGCTGAGCAGGTGCAGAGAGCTGCTGAGCTCCCAAGTCAACCAAGACTGAGGCG 301

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100

DB 302 CTGATGAGCAGACCAATGAAGAGTGAAGCCCTACAAATCGGAACGAGGAACAACGTG 361

QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120


```
|||||
Db 362 ACCCGGTACGAGAGAGACGCGGCGCATCTCCAGAGAGCTCAGACGCGCGAGGCC 421
QY 121 ArgLeng1yAlaaspmetGluaspValCysGlyArgLeuValGlnIrrArgGlyGluVal 140
Db 422 CGCGTGGCGCGGACATGGAGACGTGTGGCGCGCTGGTGTGATACCGCGCGAGGTG 481
QY 141 GlnAlaMetLeuGlyInserThrGluGluValArgLeuValAspSerHisLeuArg 160
Db 482 CAGGCCATGTGGCGCAGAGACCGAGAGCTCGGGGTGGCTCCCTCCACCTGGCC 541
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AACCTCGTAAAGCGGCTCTCCCGCGATGCCGATGACCTGAGAGAGCGCTGGCAGGTAC 601
QY 181 GlnAlaGlyAlaArgGluGluValArgGlyLeuSerAlaIleArgGluValArgLeuGly 200
Db 602 CAGGCGGGGGCGCGCAGAGGGCGCGAGCGGCTCAGCGCCATCCGCGAGCGCGTGGGG 661
QY 201 ProLeuValGluGluGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
Db 662 CCCCTGTGGAACAGGGCGCGCTGGCGCGCCACTGTGGGCTCCCTGGCGCGCCAGCCG 721
QY 221 LeuGluGluArgAlaGlnAlaIrrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
Db 722 CTACAGAGAGCGCGCCAGGCGCTGGCGCGAGCGGCTCGCGCGCGGATGAGAGATGGGC 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGluValAlaGluValArgAlaLys 260
Db 782 AGTCGGACCCGCGACCGCTGGAGCGAGGTAAAGAGAGAGAGTGGAGGGTGGCGCGCAAG 841
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaIrrpGluAlaArgLeuLys 280
Db 842 CTGGAGAGAGAGCCAGAGATAGCTGTGCAAGCGGAGGCTTCCAGAGCGCGCCCTCAAG 901
QY 281 SerTrrpHeGluProLeuValGluAspMetGlnArgGlnTrrpAlaGlyLeuValGluLys 300
Db 902 AGCTGTTCGAGCCCTGTGGAGACATGACAGCGCAGGCGCGCGCTGTGGAGGAAG 961
QY 301 ValGlnAlaAlaValGlyThrSerAlaIleProValProSerAspAsnHis 317
Db 962 GTGCAGGCTGCGGTGGGACACAGCGCGCGCTGTGTGCCAGGACATATCAC 1012

RESULT 9
US-09-880-107-2244
: Sequence 2244, Application US/09880107
: Patent No. US20020142981A1
: GENERAL INFORMATION:
: APPLICANT: Horne, Darci T.
: APPLICANT: Vockley, Joseph G.
: APPLICANT: Scherf, Uwe
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
: FILE REFERENCE: 44921-5028-WO
: CURRENT APPLICATION NUMBER: US/09/880,107
: PRIOR FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2244
: LENGTH: 1157
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20020142981A1 M12529
US-09-880-107-2244

Alignment Scores: 4.11e-136 Length: 1157
Pred. No.: 1568.00 Matches: 314
```

```
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 99.05% Mismatches: 3
Query Match: 98.68% Indels: 0
DB: 10 Gaps: 0

US-09-827-854-18 (1-317) x US-09-880-107-2244 (1-1157)

QY 1 MetLysValLeuTrrpAlaIleLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAAGTTCTGTGGGCTGCGTTGCTGTGTCACATTCTTGGCAGAGATGCCAGGCAAGGTG 121
QY 21 GluGlnAlaValGluThrGluProGluLeuArgGlnGlnIrrpGlnInser 40
Db 122 GAGCAGAGCGGTGAGACAGAGCGGAGCCGAGACTGCCACAGCAGACCGAGTGGCAGAGC 181
QY 41 GlyGlnArgTrrpGluLeuAlaLeuGlyArgPheTrrpAspTyrLeuArgTrrpValGlnThr 60
Db 182 GGCACGCGCTGGGAACTGGCACTGGGTCCCTTTGGGATTAACCTGCGCTGGGTGCAGACA 241
QY 61 LeuSerGluGlnValGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 242 CTGTCTGACAGAGTGCAGAGAGAGCTGTACGCTCCCAAGTCAACCAAGACTGAGAGGCC 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu 100
Db 302 CTGATGAGAGACATGAAAGGAGTTGAAAGGCTTACAAATCCGAACCTGAGAGAACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIrrpAla 120
Db 362 ACCCGGTAGCGAGAGAGCGCGCGCGCTGTCCAGAGCTGAGACGCGCGCAGGCC 421
QY 121 ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnIrrpArgGlyGluVal 140
Db 422 CGCGTGGCGCGGACATGAGAGACGTGTGCGCGCGCTGTGTGATACCGCGCGAGGTG 481
QY 141 GlnAlaMetLeuGlyInserThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCCATGCTCGGCGCAGAGACCGAGAGCTCGGGTGGCGCTCCCTGCCCTCCACCTGCC 541
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AACCTCGTAAAGCGGCTCTCCCGCGATGCCGATCCAGAGACTTCAGAGAGCGCTGGAGGTAC 601
QY 181 GlnAlaGlyAlaArgGluGlnIleArgGluValArgGlyLeuSerAlaIleArgGluValArgLeuGly 200
Db 602 CAGGCCGGGGCGCGCAGAGGGCGCGAGCGCGCTTACAGCGCCATCCGCGAGCGCGCTGGGG 661
QY 201 ProLeuValGluGluGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
Db 662 CCCCTGTGGAACAGGGCGCGCGTGGCGCGCACTGTGGGCTCCCTGGCGCGCCAGCCG 721
QY 221 LeuGluGluArgAlaGlnAlaIrrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
Db 722 CTACAGAGAGCGCGCCAGGCTGGGGCGAGCGGCTGGCGGGTGGAGGAGATGGGC 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGluValAlaGluValArgAlaLys 260
Db 782 AGTCGGACCCGCGACCGCTGGAGCGAGGTAAAGAGAGAGTGGAGGGTGGCGCGCAAG 841
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaIrrpGluAlaArgLeuLys 280
Db 842 CTGGAGAGAGAGCCAGAGATAGCTGTGAGCGCGAGCGGCTTCCAGGCGCGCGCTCAAG 901
QY 281 SerTrrpHeGluProLeuValGluAspMetGlnArgGlnTrrpAlaGlyLeuValGluLys 300
Db 902 AGCTGTTCGAGCCCTGTGGAGACATGACAGCGCAGATGGCGCGCGCTGTGGAGGAAG 961
QY 301 ValGlnAlaAlaValGlyThrSerAlaIleProValProSerAspAsnHis 317
Db 962 GTGCAGGCTGCGGTGGGACACAGCGCGCGCTGTGTGCCAGGACATATCAC 1012

RESULT 10
US-09-827-854-7
```



```

QY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluLeu 100
Db      301 CTGATGGACGAGACCATGAAAGGAGTTGAAGGCTTACAAATCGGAAGCTGAGAACTTC 360
QY      101 ThrProValAlaGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db      361 ACCCGGTGGCGGAGAGACGCGGCGCAAGGCTGTCCAAAGAGGTGCAAGCGGCGGCGCC 420
QY      121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db      421 CGCGTGGCGCGGACATGAGAGACGTGTGCGGCGCGCTGGTGCAGTACCGCGGCGAGGTG 480
QY      141 GlnAlaMetLeuGlyLysSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db      481 CAGGCACTGCTCGACACGAGACCGAGAGCTCGGGTGGCGCTCCCTCCACCTGCGG 540
QY      161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db      541 AAGCTCGTAAGGCGTCTCCCGCATGCCGATGACCTCGAAGAGTCCCTGCGCAGGTGAC 600
QY      181 GlnAlaGlyAlaArgGluGluGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db      601 CAGCGCGGCGCGCGGCGGCGCGGCGCGGCGCTCAGCGCCATCCGCGAGCGCTGGGG 660
QY      201 ProLeuValGluGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db      661 CCCCTGTGGAAACAGGCGCGCGTGGCGCGCGCATGTGGGTCCCTGGCGCGCGCAAGCGG 720
QY      221 LeuGlnGluArgAlaGlnAlaIleArgGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
Db      721 CTACAGAGAGGCGCGCGAGCGCTGGGCGGAGCGGCTGGCGCGGCGGATGAGAGATGGGC 780
QY      241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db      781 AGCGGACCGCGGCGCGCGCTGCGAGCGAGGTGAAGAGACAGTGGCGAGGTGCCCGCAAG 840
QY      261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db      841 CTGAGAGAGAGCGCGGAGAGATAGCTGTGAGCGCGGAGCGGCTTCCAGGCGCGCGCTCAAG 900
QY      281 SerTyrPheGluProLeuValGluAspMetGlnArgGlnTyrAlaGlyLeuValGluLys 300
Db      901 AGCTGTGTCAGAGCCCTGTGTGGAAAGCATGACAGCGGCGGCGGCGCTGTGTGGAGAG 960
QY      301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db      961 GTGACGGTGCCTGGGCGACAGCGCGCGCTGTGTCCCGAGCAATGCAC 1011

```

RESULT 12
US-09-967-013-5

```

; Sequence 5, Application US/09967013
; Patent No. US20020045840A1
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr, Vincent P
; TITLE OF INVENTION: METHOD FOR GENETIC ANALYSIS OF APOE DNA
; FILE REFERENCE: 11926-022001
; CURRENT APPLICATION NUMBER: US/09/967,013
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: 60/206,613
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 41907
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-967-013-5

```

Alignment Scores: 7,98e-116 Length: 41907
Pred. No.: 1368.50 Matches: 300
Score: 1368.50
Percent Similarity: 60.00% Conservative: 0

```

Best Local Similarity: 60.00% Mismatches: 6
Query Match: 86.12% Indels: 194
DB: 10 Gaps: 1
US-09-827-854-18 (1-317) x US-09-967-013-5 (1-41907)
QY      12 PheLeuAlaGlyCysGlnAlaLysValGluGlnAlaValGluThrGluProGluProGlu 31
Db      20316 TTCACACAGAGATGCCAGRCACAGGTGGAGCAAGCGGTGGAGACAGCGCGAGCCGAG 20375
QY      32 LeuArgGlnGlnThrLutTyrPglInseryGlnArgTyrPgluLeuAlaLeuGlyArgPhe 51
Db      20376 CTGCGCCACAGACCGAGGCGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 20435
QY      52 TyrAspTyrLeuArgTyrPglInseryGlnArgTyrPgluLeuAlaLeuGlyArgPhe 71
Db      20436 TGGGATTACCTGCGCTGGGTGGGTCAGACACTGTCTGACAGGTGACGAGGAGGAGTGTACG 20495
QY      72 SerGlnValThrGlnLysLeu 78
Db      20496 TCCACAGTACCCAGAGACAGTGAAGTGTCCCATTCCTGAGCCCTTCGTGAG 20555
QY      78 78
Db      20556 GCGCGCTATACCTCCCGAGGTCAGGTTTCATTCGCCCCCTGTGCTAAGTCTGGGGGG 20615
QY      78 78
Db      20616 CTTGGGTCTGTGCTGTGTCTAGCTTCCTCTCCATTTCTGACTCCTGCTTAACTCTTC 20675
QY      78 78
Db      20676 TGGAAATTCCTCTCTCAGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCACT 20735
QY      78 78
Db      20736 CGTCTGCTCTGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 20795
QY      78 78
Db      20796 CTCACGTGTGTGCCAGGCTGTCTTGAACCTTGTGGGCTCAAGCATCTCCCGCTCGG 20855
QY      78 78
Db      20856 CCTCCCAAGATGCTGGGATTAGAGCATAGCACCTTGCCGCGCTCTAGTCTCTTCT 20915
QY      78 78
Db      20916 TCGTCTGTGCTCTGCTCTGCTCTGATCTGCTCTGTGATCTGCTCTGTCTCTCTCTG 20975
QY      78 78
Db      20976 GCCTCTGCCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 21035
QY      78 78
Db      21036 CCCCATCCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 21095
QY      79 -ArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGln 98
Db      21096 CAGGCGCGGTGATGAGACGACCATGAAGAGTGAAGGCTTCAAAATCGGAAGTGAAGAGA 21155
QY      98 uGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAla 118
Db      21156 ACAACTGTACCCCGGTGGCGGAGAGAGCGGCGGCGGCTGTCCAAAGACTCTGAGGCGGC 21215
QY      118 acGlnAlaArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGln 138
Db      21216 GCAAGCGCGCGGTGGCGGCGGACATGAGAGACAGTGTGCGGCGCGCTGTGTGACGTACGCGG 21275
QY      138 yGluValGlnAlaMetLeuGlyLysSerThrGluLeuArgValArgLeuAlaSerHis 158
Db      21276 CGAGGTGCGAGGCATCTGCGCGAGAGCACCGAGAGCTGCGGCTGCGCTGCGCTGCCA 21335

```


Db 396 ATGAGGCTTCGTGGCGCTGGCTGTGCACATTCCTGGCAGAGCCAGCCAAAGTG 337
QY 21 GUGLAlaValgluThrgluProgluLeuAArglGluThrgluTrpGlnSer 40
Db 336 GACCAAGCGGTGGAGACAGCCGGAGCCNCCAGCTGCGCAGAGACCCAGTGGCAGAGC 277
QY 41 GUGLArgTrpGluLeuAlaLeuGlyAArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 276 GGCACGCGTGGGAATGGACACTGGCTCGCTTTGGGATTACCTGGCGTGGGTCAGACA 217
QY 61 LeuSerGluGlnValGlnGluLeuLeu-SerSerGlnValThrGlnGluLeuArgAl 80
Db 216 CTGTCTGACAGGTGACAGAGGAGGAGGCTGCTCCAGCTCCAGCTCAGCCAGAACTGAGGGC 157
QY 80 AleuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGln 100
Db 156 GCGATGAGCAGAGACCATGAAGGCTTGAAGGCTTCAAAATCGGAATCGAGAACAACT 97
QY 100 urThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAl 120
Db 96 GACCCGGGTGGCGGAGAGACGCGGCGCTGTCCAAAGAGACTGACAGCGCGGAGAGN 37
QY 120 aArgLeuGlyAlaAspMetGluAspVal 129
Db 36 CCGGCTGGGCGCGACATGAGAGACGTG 9
RESULT 15
US-09-880-107-2491/c
; Sequence 2491, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2491
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 N33009
; LOCATION: (1)..(478)
; OTHER INFORMATION: n = a o r c o r g o r t
US-09-880-107-2491
Alignment Scores:
Pred. No.: 1.27e-49 Length: 478
Score: 625.00 Matches: 126
Percent Similarity: 97.69% Conserved: 1
Best Local Similarity: 96.92% Mismatches: 2
Query Match: 39.33% Indels: 1
DB: 10 Gaps: 0
US-09-827-854-18 (1-317) x US-09-880-107-2491 (1-478)
QY 1 MetLysValLeuTrpAlaLeuValTrpPheLeuAlaGlyCysGlnAlaLysVal 20
Db 396 ATGAAAGCTTCTGGGCTCGCTTGTGCTGCTGACATTCCTGGCAGATGCCAGCAAGTG 337
QY 21 GUGLAlaValgluThrgluProgluLeuAArglGlnGluThrgluTrpGlnSer 40
Db 336 GACCAAGCGGTGGAGACAGCCGGAGCCNCCAGCTGCGCAGAGACCCAGTGGCAGAGC 277

QY 41 GUGLArgTrpGluLeuAlaLeuGlyAArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 276 GGCACGCGTGGGAATGGACACTGGCTCGCTTTGGGATTACCTGGCGTGGGTCAGACA 217
QY 61 LeuSerGluGlnValGlnGluLeuLeu-SerSerGlnValThrGlnGluLeuArgAl 80
Db 216 CTGTCTGACAGGTGACAGAGGAGGAGGCTGCTCCAGCTCCAGCTCAGCCAGAACTGAGGGC 157
QY 80 AleuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGln 100
Db 156 GCGATGAGCAGAGACCATGAAGGCTTGAAGGCTTCAAAATCGGAATCGAGAACAACT 97
QY 100 urThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAl 120
Db 96 GACCCGGGTGGCGGAGAGACGCGGCGCTGTCCAAAGAGACTGACAGCGCGGAGAGN 37
QY 120 aArgLeuGlyAlaAspMetGluAspVal 129
Db 36 CCGGCTGGGCGCGACATGAGAGACGTG 9

Search completed: March 15, 2003, 03:04:53
.. Job time : 81.8753 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:09:47 ; Search time 1289.69 seconds

(without alignments)
3980.771 Million cell updates/sec

Title: US-09-827-854-18

Perfect score: 1589

Sequence: 1 MKVIMALLVTFIACGQAKV.....VEKVOAAVTSAPVPSDNH 317

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus.p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool/US09827854/runat_11032003_101610_27495/app.query.fasta_1.3576
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=100 -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854.ecn1.1.882c@runat_11032003_101610_27495 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MAP -LARGEOUTER -NEG_SCORES=0 -WAIT -LONGLOG -DEV=TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1:	em_estba:*
2:	em_esthum:*
3:	em_estlin:*
4:	em_estlun:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_hlc:*
9:	gb_estl:*
10:	gb_est2:*
11:	gb_hlc:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	gb_gss:*
18:	em_gss_hum:*
19:	em_gss_liv:*
20:	em_gss_pln:*
21:	em_gss_vrt:*
22:	em_gss_fun:*
23:	em_gss_mam:*
24:	em_gss_mus:*
25:	em_gss_other:*
26:	em_gss_pro:*
27:	em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1282.5	80.7	1027	13	B1670367	B1670367 603292738
2	1246	78.4	842	12	B6763371	B6763371 602735433
3	1240	78.0	933	13	B1668318	B1668318 603295681
4	1212	76.3	938	12	B6761746	B6761746 602717942
5	1209.5	76.1	942	13	B1600906	B1600906 6032449241
6	1195	75.2	800	13	BM042094	BM042094 603615713
7	1194.5	75.2	927	12	B6472299	B6472299 602513830
8	1191.5	75.0	922	12	B1597743	B1597743 603248609
9	1182.5	74.4	817	12	B6774871	B6774871 602649975
10	1181	74.3	811	13	B1600563	B1600563 603244936
11	1167.5	73.5	845	12	B6829472	B6829472 602763768
12	1144	72.0	790	12	B6707147	B6707147 602670283
13	1141	71.8	919	13	B1551475	B1551475 603194314
14	1131	71.2	757	13	BM042228	BM042228 603616186
15	1121.5	70.6	907	12	B6706129	B6706129 602669093
16	1120	70.5	706	14	BM728696	BM728696 UI-E-EJO-
17	1120	70.5	741	12	B6762924	B6762924 602735153
18	1118	70.4	1110	11	AK010261	AK010261 Mus. muscu
19	1116	70.2	804	12	B6702752	B6702752 602684616
20	1115	70.2	834	13	BM042676	BM042676 603616028
21	1112.5	70.0	808	13	B1668329	B1668329 603295692
22	1109	69.8	855	13	B1161362	B1161362 602865769
23	1106	69.6	803	13	B1670350	B1670350 603292721
24	1103	69.4	782	12	B6716776	B6716776 602678182
25	1103	69.4	812	13	B1601551	B1601551 603250760
26	1103	69.4	965	14	B6677266	B6677266 AGENCOURT
27	1100	69.2	757	13	B1603658	B1603658 603251982
28	1097.5	69.1	797	12	B6715366	B6715366 602675503
29	1097	69.0	846	13	B1159757	B1159757 602863509
30	1096	69.0	796	13	BM042153	BM042153 603615790
31	1091	68.7	757	12	B6707750	B6707750 602671277
32	1091	68.7	794	13	B1601279	B1601279 603245010
33	1091	68.7	954	12	BF667543	BF667543 602287404
34	1089	68.5	790	13	B1551066	B1551066 603196425
35	1088	68.5	790	13	B1551811	B1551811 603192716
36	1084.5	68.3	798	12	B6708414	B6708414 602672374
37	1083.5	68.2	891	13	B1549292	B1549292 603189930
38	1077.5	67.8	802	13	B1458355	B1458355 603198705
39	1076	67.7	1100	14	BM914382	BM914382 AGENCOURT
40	1075.5	67.7	914	13	B1603523	B1603523 603244424
41	1075	67.7	748	13	B1553085	B1553085 603196058
42	1074.5	67.6	930	12	BF667857	BF667857 602287176
43	1070.5	67.4	812	12	B6769968	B6769968 602739854
44	1070	67.3	688	13	B1544886	B1544886 603242328
45	1069.5	67.3	803	12	B6709360	B6709360 602673385

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
B1670367	603292738F1	NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5312024 5'	B1670367	1	GI:15584600	human.	Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	1 (bases 1 to 1027)	NIH-MGC http://imgc.ncl.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA library Preparation: Michael J. Brownstein (NHGRI), Shiroh
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LRAM11790 row: 1 column: 09
 High quality sequence: 100%

Location/Qualifiers
1. .1027

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5312024"
/clone_id="NH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pBluescriptp (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagc
); Oligo-dt primed using primer 5'-tttttttttttttttVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROP 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NH_MGC Library."

```

Pred. No.:	2,15e-124	Length:	102
Score:	1282.50	Matches:	291
Percent Similarity:	90.55%	Conservative:	6
Best Local Similarity:	88.75%	Mismatches:	20
Query Match:	80.71%	Indels:	12
DB:	13	Gaps:	2

US-09-827-854-18 (1-317) x BI670367 (1-1027)

QY	1	MettYsValLeuThrPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	26	ATTAAGGTTCTGTGGGCTGGCTGGCTGCATCTCTGGCAGAGATCCAGGCAAGGGT	85
QY	21	GIuGlnAlaValAlaGluThrGlnuProGluuProGluLeuArgGlnGlnThrgluThrpGlnSer	40
Db	86	GACCAAGCGGTGGAGACAGACGCGGAGCGCCAGTGGCGCAGACAGACAGTGGCGAGAC	144
QY	41	GlyGlnArgTrpGluLeuAlaLeuGlnLysArgPheTrpAspTyrLeuArgTrpValGlnThr	60
Db	146	GGCCACCGCTGGGAACTGGCACTGGGTGCGCTTTGGATTACCTGGCTGGGTCAGACA	204
QY	61	LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnLeuLeuArgAla	80
Db	206	CTGTCTGACACAGGTCCAGAGAGACTGCTCAGCTCCAGATCCACGAGAACTGAGGGCG	264
QY	81	LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu	100
Db	266	CTGATGTGACGAGACCATTAAGAGACTTGAAGGCTTCAAAATCGAACTGGAGGAAACAACGTG	324
QY	101	ThrProValAlaGluGluThrArgAlaArgPheSerLysGlnLeuGlnAlaAlaGlnAla	120
Db	326	ACCCGCGTGGCGAGAGAACCGCGGGCACGCGTGTCAAAAGAGCTGCAGCGCGGCGCAGGCC	384
QY	121	ArgLeuGlnAlaAspMetGluAspValCysGlnYargLeuValGlnTyrArgGlyGluVal	140
Db	386	CGCGTGGCGCGGACATGAGAGACGTGTGGCGGCCCTGGTGCATGACCGCGGCGAGGTG	444
QY	141	GlnAlaMetLeuGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg	160

Db	446	CAGGCCATGCTCTGGCCGACAGCACCGGAGGCGTGGGGTGGCTGCTCCCTCCACCTGGCC	505
OY	161	LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValI	180
Db	506	AAGCGCGCTAAGCCGGCTCTCCCGCATGCCATGTACCTTCACAAAGCGCTGGCAGGTAC	565
OY	181	GlnAlaGlyAlaArgGlnGlyAlaArgGlyLeuSerAlaIleArgGlnArgLeu-Gl	200
Db	566	CAGGCGGGGGCCGCGAGGGCCGACGCGGGCTTCAGCGCATCCGCGAGCGCTGGGG	625
OY	200	ProLeuValGlnGlnGlyArgValArgAlaIleThrValGlySerLeuAlaGlyInP	220
Db	626	CCCCCTGGTGGACAGGCTCCGCTGGCGGGCCGCCACTGTGGGCTCCCTGGCCGACGC	685
OY	220	OLEuGlnLuarArgAlaGlnAlaIleTrpGlyGluArg-LeuArgAlaArgMetGluGln-Met	239
Db	686	GCTACAGGAGCGGGCCAAAGGCTGGGGGAGCGGGCTTGGCGCGCGCATGAGAGAGAG	745
OY	240	GlySerArg-ThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlu-ValArg	259
Db	746	GGCAGCCCGGAGCCCGCCGACCGCTGGACCGAGGTGAAGGACGAGTGGCGGAAGTCCGCG	805
OY	259	IaLysLeu---GluGlnGlnAlaGlnGln-IleArgLeuGln-AlaGlnAlaPheGlnI	277
Db	806	CCCCAGCTGGACGACGACCGAGGCCAGCAAGATAGCGCTTCAGGCGGAGGCTTCTAGGC	865
OY	277	ArgLeuLysSerTrpPheGluPro-----LeuValGluAspMetGluArgGlnTrpAl	295
Db	866	CCGGCTCTCAGAGCTGCTTCGAACCCCGTGGCTTGGACACACATGGCGGGCGCACGTGGCG	925
OY	295	AGlyLeuValGlnLysVal-GlnAlaAlaValGlyThrSerAlaAlaProValProSer	315
Db	926	CGGCTGCTGTAAGAAAGGTGTCAGGCTGCCGGGGGACACACCCCGCTGTGTCCAGC-G	984
OY	315	spasnHis 317	
Db	985	ACAATCAT 992	

LOCUS

DEFINITION 602335433F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4860585 5', mRNA sequence.

KEYWORDS

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHOR

TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 Clone Sequencing by: Incyte Genomics, Inc.
 DNA Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:

nos

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4860585"
/clone_lib="NIH MGC 49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"

```


QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySer-LeuAlaGlyGlnPro 220
|||||
Db 657 CCCCTGTTGACACAGGCGCGCGCGCGCCACTGTGGCTTCCTGGCCGCGCAGGCC 716
QY 220 OleuGlnGluArgAlaGlnAlaTrp-GlyGluArg-LeuArgAlaArgMetGlu--Glu 238
|||||
Db 717 GCTACAGAGACCGGCAT 776
QY 239 MetGlySerArgThr-Arg-AspArgLeuAspGluValLys-GluGlnValAla---Glu 256
|||||
Db 777 TGGCGACGCCGACCG 836
QY 257 ValArgAlaLysLeuGlnGluGln-GlnAlaGlnGlnLeuArgLeuGlnAla--GlnAlaPhe 275
|||||
Db 837 TTGGGCGCCCAAGCTGGAGAGACAGCGCCAGGAGATACGCTCGCAGCGCGCGCGCGCGCGCG 896
QY 276 GlnAlaArgLeuLysSerTrpPheGluProLeuVal 287
|||||
Db 897 CCAGCGCGCTCAAAAGCTGTCTGAGAGCGCTGTGT 932

RESULT 5
Bi600906 942 bp mRNA linear EST 07-SEP-2001
LOCUS 603249241F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5301010 5',
DEFINITION mRNA sequence.
ACCESSION Bi600906
VERSION Bi600906.1 GI:15493845
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 942)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgrabbs-r@mail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshitaki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLA011762 row: a column: 11
High quality sequence stop: 762.
Location/Qualifiers
1. 942
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5301010"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site:1: BamHI; Site:2: SalI-XhoI (qtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTAA-3',
size-selected for average insert size 2.3 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI, National
Institutes of Health). Note: this is a NIH-MGC library."

BASE COUNT 177 a 284 c 361 g 120 t
ORIGIN

Alignment Scores: 8,94e-117 Length: 942
Pred. No.: 1209.50 Matches: 265
Score: 93.13% Conservative: 6
Percent Similarity:

Best Local Similarity: 91.07% Mismatches: 14
Query Match: 76.12% Indels: 8
DB: 13 Gaps: 1
US-09-827-854-18 (1-317) x Bi600906 (1-942)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
Db 73 ATGAGAGTTCTGTGGGTGGCTGTGCTGTCACATTCCTGTCAGAGATCCAGGCAAGGTG 132
QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
|||||
Db 133 GAGCAAGCGGTGAGACAGAGCCGAGCCGAGCTGCGCCAGCAGACCCGAGTGGCAGAGC 192
QY 41 GlyGlnArgTrpGluLeuAlaLeuGluArgPheTrpAspPyrLeuArgTrpValGlnThr 60
|||||
Db 193 GGCACGCGCTGGAACTGGGACGCGGCTGCTTTGGGATTAACCTGCGCTGGGTGCAGACA 252
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
Db 253 CTGCTCAGCAGGTGCGAGAGAGAGCTGCTCAGCTCCAGGTCACCAAGAACTGAGGCGC 312
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu 100
|||||
Db 313 CTGATGGACGAGACCATGAGAGATTGAAGGCTTACAAATCGGAATCGAGAAACAACCTG 372
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
|||||
Db 373 ACCCGGTGGCGAGGAGAGCGCGGCGCGCTGCTCCAAAGAGCTGCGGCGCGCAGGCC 432
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
|||||
Db 433 CGCTGGCGCGCGGACATGAGAGAGCTGTGGCGCGCTGTGTCAGATACCGCGGAGGTG 492
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 493 CAGCGCATGCTCGCGCGAGACACCGAGAGCTCGCGCTCGCGCTCCACCTCGCGC 552
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
Db 553 AAGCTGCGTAAAGGCGCTCTCCGCGATGCGGATGCTGCGAAGCGCTGCGCAGTTAC 612
QY 181 GlnAlaGlyAlaArgGlnGlnAlaGluArgGlyLeuSerAlaLeuArgGluArgGly 200
|||||
Db 613 CAGGCGCGCGCGCGCGAGGCGCGGAGCGCGGCTCAGCGCCATCCGCGAGCGCTGGGG 672
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
|||||
Db 673 CCCCTGTGGACAGGCG 731
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAla-ArgMetGlu-GluMetG 240
|||||
Db 732 CTACAGAGAGGCG 791
QY 240 LysSerArgThrArgAspArgLeuAsp-GluValLys-GluGlnValAlaGluVal-ArgA 259
|||||
Db 792 GCAAGCGGACCGCGAGCGCGCTGAGCGAGTGAAGCGACAGAGTGGCGAATGGCGCG 851
QY 259 LysLeuGlnGluGlnAla--GlnGlnLeuArgLeuGlnAlaGluAlaPheGlnAla 278
|||||
Db 852 CCAACTTGGAGAAACAGCGCGCCAGCAATACGCTTGCAGCGCGAGCGCTTCAGCGCC 911
QY 278 rLeuLysSerTrpPheGluPro 285
|||||
Db 912 GCCTCAAAAC-TGGTTGACCCCT 933

RESULT 6
BM042094 800 bp mRNA linear EST 07-NOV-2001
LOCUS 603615713F1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5420618 5',
DEFINITION mRNA sequence.
ACCESSION BM042094
VERSION BM042094.1 GI:16771361

KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 800)
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLCM1875 row: 1 column: 03
High quality sequence: 792.

FEATURES	Location/Qualifiers
source	1. .800

BASE COUNT 146 a 239 c 315 g 100 t
ORIGIN
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5420618"
/clone_1ib="NIH_MGC_112"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCGAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

Alignment Scores:	
Pred. No.:	2,36e-115
Score:	1195.00
Percent Similarity:	98.80%
Best Local Similarity:	97.20%
Query Match:	75,20%
DB:	13
Length:	800
Matches:	243
Conservative:	4
Mismatches:	3
Indels:	1
Gaps:	0

US-09-827-854-18 (1-317) x BM042094 (1-800)

QY	1	MeltyValLeuTPtPAlaIaLeuLeuValThrPheLeuAlaIaCYSGAlaIaLysAla	20
Db	50	ATGAAGGTTCTGTGGCGTGGCTGCTGTGCATTTCTCGGAGGATGCGACGCCAAGTG	100
QY	21	GIuGlnAlaValaIGluThrGluProGluProGluLeuArgGlnGlnThrGluTPGInSer	40
Db	110	GAGCAAGCGGTGGAGACAGACCGGAGCCGAGGTGGCCGACGACCGAGTGGCAAGC	160
QY	41	GlyGlnArgTPtPGluLeuAlaIaLeuGlyAArgPheTPtPArgTyrLeuArgTyrPAlaThr	60
Db	170	GGCCAGCGCTGGAACTGGCACTGGGTGGCTTTTGGGATTACCTGGCGTGGGTGCAGCA	220
QY	61	LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80
Db	230	CTGTCTAGACGGGTGACGAGGAGACTCTCAAGTCCACAGGTACCCAGGAATGAGGGCG	280
QY	81	LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGluLeu	100
Db	290	CTGATGACGAGACCATGAAGAGATTGAAGGCTTACAATGGCAACTGGAGAACCACTG	340
QY	101	ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaIaGlnAla	120
Db	350	ACCCCGGTGGGGAGGAGACCGGGCACGGCTGTTCACAAAGAGGTGCAGGGCGCCAGGCC	400

Qy	121	ArgleugllyAlaAspMetGluAspAlaCysGlyArgLeuValGlnTyrArgGlyGluVal	140
Db	410	CQGGCTGGGGCCGCAATGAGAGCACTGTGCGGCCCGCTGTGCTACGTACGGCGCAGGTG	469
Qy	141	GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	470	CAGGCCATGCTCTGGCCAGAGCACCCAGAGAGCTGGGGGTGCGCTCGCTCCACCTGGCC	529
Qy	161	LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	530	AAGCTGCCTAAACGGGCTCTCCGCAATGCCGATGACCTGCACAAACCGCTGGCACTGTAC	589
Qy	181	GlnAlaGlyAlaArgGlnGlyValGlnArgGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly	200
Db	590	CAGGCCGGGGCCCCCGGAGGGCCGCCAGCGCGGCTCTCAGCGCCATCCGCGAGCGCTTGGGG	649
Qy	201	ProLeuValGlnGlnGlyArgValAlaArgAlaAlaThrValGlySerLeuAlaGlyGlnPro	220
Db	650	CCCTGTGGTGAACAGGGCCGCGTGGGGGCGACACTGTGGGCTCCCTGGCCGCGCAGCGG	709
Qy	221	LeuGlnGlnAlaArgAlaGlnAlaAlaTrpGlyGlnArgLeuArgAlaArgMetGlnIleMetGly	240
Db	710	CTACAGGAGCGGGCCCGACGGCTGTGGGGCGCAGCGGTG-CGCGCGCGGATGAGAGAGATGGGC	768
Qy	241	SerArgThrArgAspArgLeuAspGluVal	250
Db	769	ACGGCGGACCGGGAACGCTGGACACGTG 798	

LOCUS B6472229 927 bp mRNA linear EST 21-MAR-2001
DEFINITION 60251830P1 NIH_MGC_16 Homo sapiens CDNA clone IMAGE:4645759 5',
mRNA sequence.
B6472229
ACCESSION B6472229.1 GI:13404485
VERSION
KEYWORDS EST.
SOURCE human.

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 927)
TITLE	NIH-MGC http://mgc.nci.nih.gov/
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999)
	Contact: Robert Strausberg, Ph.D.

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (tLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MCC clone distribution information can be
found through the I.M.A.G.E. Consortium/tLNL at: image.llnl.gov
Plate: LLCM419 Row: K Column: 08
High quality sequence stop: 846.

FEATURES
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4645759"
/clone_1lb="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pORB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

```


RESULT 10
 BI600563
 LOCUS 811 bp mRNA linear EST 07-SP-2001
 DEFINITION 60324936F1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:5287329 5',
 mRNA sequence.
 ACCESSION BI600563
 VERSION BI600563
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 811)
 NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgrabbs-remail.nih.gov
 Tissue Procurement: Miklos Palokvits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shuraki
 Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHAM1726 row: g column: 10
 High quality sequence stop: 783.
 Location/Qualifiers
 1..811
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5287329"
 /clone_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescriptK1 (modified
 pBluescript K1); Site: 1: BamHI; Site: 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
 size-selected for average insert size 2.3 kb and
 normalized to 10^5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapped method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH-MGC library."
 BASE COUNT 148 a 245 c 316 g 102 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 7.12e-114 Length: 811
 Score: 1181.00 Matches: 240
 Percent Similarity: 97.978 Conservative: 1
 Best Local Similarity: 97.568 Mismatches: 4
 Query Match: 74.328 Indels: 1
 DB: 13 Gaps: 0
 US-09-827-854-18 (1-317) x BI600563 (1-811)
 QY 1 MettlyValleutrpAlaalaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||||||
 DB 74 ATGAAGTTCTGTGGCTGCTGTGGTGCATTCCTCGCAGATGCCAGGCCAAGTGC 133
 |||||||
 QY 21 GtuglAlaValGlturhrgtugProGluProGluLeuAArgInGlnInhrGlturtpGlnSer 40
 |||||||
 DB 134 GACCAAGCGGTGAGACAGAGCCGAGCCGAGCTCGCCAGCAGACCGAGTGGCAGAGC 193
 |||||||
 QY 41 GtlyGlnArGtrpLutleuAlaLeuGlyArGpPheTrpAspTrpLeuArGtrpValGlnThr 60
 |||||||
 DB 194 GCGCAGCGCTGGGAACTGCACACTGGGTTCGTTTGGGATTACTGCGCTGGGTCGACACA 253
 |||||||
 QY 61 LeuSerGluGlnValGlnGluGluLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||||||
 DB 254 CTGTCTAGACAGGTGACGAGAGCTGTCTACCTCCAGGTACACGAGAACTGAGGGCG 313

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 |||||||
 DB 314 CTGATGACGACAGACCATGAAGAGCTTGAGAGCCTTACAAATCGGAACTGGAGAACAACTG 373
 |||||||
 QY 101 ThrProValAlaGluGluThrArGAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
 |||||||
 DB 374 ACCCGGTGGCGGAGGAGACGCGGGCAGCGCTCTCCAAAGAGCTGCAGGCGGGCGAGGCC 433
 |||||||
 QY 121 ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
 |||||||
 DB 434 CGGCTGGCGCGGACATGGAGACGTGTGGCGCGCTGTGTGCATGCCCGCGAGGTG 493
 |||||||
 QY 141 GtAlaMetLeuGlyGlnSerThrGluGluLeuArGValArgLeuAlaSerHisLeuArg 160
 |||||||
 DB 494 CAGGCCATGCTCGGCCACAGACACCGAGAGACTCGGGTGGCTCGCTCCACTCGCGC 553
 |||||||
 QY 161 LysLeuCysLysArgLeuLeuArGAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||||||
 DB 554 AAGCTGCTGAAGGGGCTCTCCCGCATGCGCATGACCTGCAGAAAGCCCTGGCAGTGTAC 613
 |||||||
 QY 181 GtAlaAGtlyAlaArgGluGluAlaGluArgGlyLeuSerValIleArgGluArgLeu 200
 |||||||
 DB 614 CAGCGCGGGCGCGCGAGGGCGCGAGCGGGCTCAGCCCATCCGCGAGCGCTGGGG 673
 |||||||
 QY 200 yProLeuValGluGlnGlyArGValArGAlaAlaThrValGlySerLeuAlaGlnPr 220
 |||||||
 DB 674 CCCCTTGTTGAGACAGAGCGCGCTCGCGCGCCACTGTGGCTCTCCGCGACAGCGCC 733
 |||||||
 QY 220 OleuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluLumetG 240
 |||||||
 DB 734 GTACAGAGACGCGGCCAGAGCTGTGGCGAGCGGCTGCGCGCGATGTAGGCAATGGG 793
 |||||||
 QY 240 ySerArgThrArgAsp 245
 |||||||
 DB 794 CAGCCGGAACCGCGAA 809
 |||||||
 RESULT 11
 BG829472 845 bp mRNA linear EST 22-MAY-2001
 LOCUS BG829472
 DEFINITION 602763768F1 NIH_MGC_42 Homo sapiens CDNA clone IMAGE:489112 5',
 mRNA sequence.
 ACCESSION BG829472
 VERSION BG829472.1 GI:14177059
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 845)
 NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgrabbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHAM1790 row: g column: 17
 High quality sequence stop: 829.
 Location/Qualifiers
 1..845
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:489112"
 /clone_lib="NIH_MGC_42"
 /tissue_type="epithelioid carcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: pancreas; Vector: pOMB7; Site: 1: XhoI;

Site_2: EcoRI. cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. 1"

BASE COUNT 153 a 244 c 342 g 106 t
ORIGIN

Alignment Scores:

Pred. No.: 1.98e-112 Length: 845
Score: 1167.50 Matches: 252
Percent Similarity: 95.85% Conservative: 2
Best Local Similarity: 95.09% Mismatches: 8
Query Match: 73.47% Indels: 5
DB: 12 Gaps: 2

US-09-827-854-18 (1-317) x BG829472 (1-845)

QY 1 MetlyValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
Db 57 ATGAAGGTTCTGTGGCGTGGCTGTGTCATTCCTGGCAGAGATGCCAGCAAGTG 116
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 117 GAGCAAGCGGTGGAGACAGAGCCGAGCGCCAGCCAGACCGAGATGGCAGAGC 176
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 177 GGCACAGCCCTGGAGACTGGCACTGGGTGCTTTTGGAATTACCTGGCTGGATGACAGA 236
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 237 CTGTGAGCAGGTGGCAGAGAGAGCTGCTCAGCTCCAGCTCCAGCAAGACTGAGGCG 296
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
Db 297 CTGATGGACGAGACCATGAAAGAGTTGAAGCCTCAAAATCGGAAGCTGGAGAACACTG 356
QY 101 ThrProValAlaGluGlnGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 357 ACCCGGTGGCGGAGAGACCGCGGACGCTGTCTCAAGAGAGCTGACAGCGCGCAGGCC 416
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 417 CGGCTGGCGCGGACATGAGAGAGCTGTGGCGCGCTGTGAGTACCGCGGAGAGTG 476
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 477 CAGGCCATGCTCGCCCAAGCAGCAGAGAGCTGGCGGCTCCCTCCACCTCGCGC 536
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 537 AAGCTGCGTAAGCGGCTCTCTCCGCGATGCCATACCTGCAAGAGCGCTTGGCAGTATG 596
QY 181 GlnAlaGlnAlaArgGluGlnGlyAlaArgGlyLeuSerAlaAlaArgGluArgLeuGly 200
Db 597 CAGGCCGCGCGCGGAGAGCGCGCGGCGCTCAGCGCGCATTCGCGAGCGCGCTGGGG 656
QY 201 ProLeuValGluGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 657 CCCCTGGTGGAAAGAGCGCGCGCGCGCGC -ACTGTGGGCTCTCCCTTGGCGGAG -CCG 714
QY 221 LeuGlnGlnAlaArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGlnGlnMetGly 240
Db 715 CTACAGAGAGCGCGCAGGCTG ---GGCGAGCGGCTGCGCGCGGAGTGAAGAGATGGCG 771
QY 241 SerArgTrpArgAspArgLeuAspGluValLysGluGlnGlnAlaGluValAlaGlnAla 260
Db 772 AGCGGAGACCGCGA ---GGCGTGGAGAGGTGAAGAGAGAGTGGCGGAGTGGCGGCCAA 828

QY 260 sleuGluGluGln 264
Db 829 GCTTGGAGAGCAG 841

RESULT 12

LOCUS BG707147

DEFINITION BG707147 790 bp mRNA linear EST 07-MAY-2001
602670283F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4793084 5',

ACCESSION BG707147

VERSION BG707147.1 GI:13983201

SOURCE

human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE NIH-MGC http://mhc.nci.nih.gov/
1 (bases 1 to 790)

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgaaps-femail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LHAM0672 row: e column: 21
High quality sequence stop: 781.

FEATURES

source

Location/Qualifiers

1..790

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4793084"

/clone_1lb="NIH_MGC_96"

/tissue_type="hypothalamus"

/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site: 1: BamHI; Site 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTAA-3',
normalized for average insert size 2.3 kb and
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 143 a 234 c 311 g 102 t
ORIGIN

Alignment Scores:

Pred. No.: 5.3e-110 Length: 790
Score: 1144.00 Matches: 235
Percent Similarity: 98.33% Conservative: 0
Best Local Similarity: 98.33% Mismatches: 4
Query Match: 71.99% Indels: 1
DB: 12 Gaps: 0

US-09-827-854-18 (1-317) x BG707147 (1-790)

QY 1 MetlyValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
Db 73 ATGAAGGTTCTGTGGCGTGGCTGTGTCATTCCTGGCAGAGATGCCAGCAAGTG 132
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 133 GAGCAAGCGGTGGAGAGACCGAGCCGAGCTGGCCAGACGAGCGATGGCAGAGC 192
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 193 GGCACAGCGCTGGAGACTGGCACTGGGTGCTTTTGGGATTAACCTGGGTGACAGACA 252

QY 61 LeuSerGluGlnValGlnGluGluLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
Db 253 CTGTCGAGAGAGGTGACGAGGAGGAGCTGCTCCAGCTCCAGGTACCAGAGAGTGGAGGGG 312
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
|||||
Db 313 CTGATGAGAGACCATGTAAGGAGTTGAAGGCTTACAAATCCGAACTGGAGGAACTG 372
QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
|||||
Db 373 ACCCGGTGGCGAGGAGAGCGCGGCGCTCCAGAGAGCTGCAGCGGCGCGAGGCC 432
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
|||||
Db 433 CGGCTGGGGCGGACATGAGAGACGCTGCGCGCCCTGGTGCAGTACCGCGCGAGGTG 492
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 493 CAGGCAATGCTCGCGCAGAGCACCGAGAGCTCGGGTGGCTCGCTCCCTCCACTGGCG 552
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
Db 553 AACCTCGTAAGCGGCTCTCCCGCATGCCGATGACCTGCAGAAAGCCCTGGCAGGTAC 612
QY 181 GlnAlaGlyAlaArgGlnGlnArgValArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||
Db 613 CAGCGCGGGGCCCGCAGAGGCCCGCGGCTCTCAGCGCCCATCCGCGAGCCCTGGGG 672
QY 201 ProLeuValGluGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
|||||
Db 673 CCCCTGTGGAACAGGGCGCGCTGCGGGCGCCACTGTGGGCTCCCTGGCGACAGCCG 732
QY 221 LeuGlnGluArgAlaGlnAlaThrGlyGluArgLeuArgAlaArgMetGluGlnLeu 239
|||||
Db 733 CTAAGAGAGCGGGCCAGGCTGGGGCGAGCGGTG-CGCGCGCGGATGAGAGAGATG 788
RESULT 13
BI551475 919 bp mRNA linear EST 05-SEP-2001
LOCUS 603194314P1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5274003 5',
DEFINITION mRNA sequence.
ACCESSION BI551475
VERSION BI551475.1 GI:15438787
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/
AUTHORS 1 (bases 1 to 919)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM11691 row: 1 column: 04
High quality sequence stop: 812.
Location/Qualifiers
1..919
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5274003"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dr primed using primer 5'-TTTGTGTGTGTGTGTGTGT-3',
size-selected for average insert size 2.5 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 172 a 270 c 363 g 114 t
ORIGIN
Alignment Scores:
Pred. No.: 1.37e-109 Length: 919
Score: 1141.00 Matches: 246
Percent Similarity: 93.26% Conservative: 17
Best Local Similarity: 92.13% Mismatches: 17
Query Match: 71.81% Indels: 4
DB: 13 Gaps: 0
US-09-827-854-18 (1-317) x BI551475 (1-919)
QY 1 MetLysValLeuThrAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
Db 73 ATGAAGTTCTGTGGGCTCGCTGCTGTGTCACATTCCTGGCAGAGATGCCAGGCAAGTGC 132
QY 21 GlnGlnAlaValGlnThrGluProGluProGluLeuArgGlnGlnThrGluTyrPglIns 40
|||||
Db 133 GAGCAAGCGGTGGAGACAGAGCCGAGCCGACCTCGCCAGAGAGACAGAGTGGCGAGAGC 192
QY 41 GlyLysArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTyrPvalGlnThr 60
|||||
Db 193 GGCACGCTGGGAACTGCGACATGGCTGCTTTTGGGATTACTGGCTGGGTCGAGACA 252
QY 61 LeuSerGluGlnValGlnGluGluLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
Db 253 CTGTCGAGAGAGGTGACGAGAGAGTGTGCTCAGCTCCAGGTACCCAGGAATGAGGGCG 312
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
|||||
Db 313 CTGATGAGAGACCATGTAAGGAGTTGAAGGCTTACAAATCCGAACTGGAGGAACTG 372
QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
|||||
Db 373 ACCCGGTGGCGAGAGAGACGCGCGGCTCTCCAGAGAGCTGCAGCGGCGCGAGGCC 432
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
|||||
Db 433 CGGCTGGGGCGGACATGAGAGACGCTGCGCGCCCTGGTGCAGTACCGCGCGAGGTG 492
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 493 CAGGCGGAGGCGGATGAGGAGCGCGGCTCTCAGCGCCATCCGCGAGCCCTGGGG 672
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
Db 553 AACCTCGTAAGCGGCTCTCCCGCATGCCGATGACCTGCAGAAAGCCCTGGCAGGTAC 612
QY 181 GlnAlaGlyAlaArgGlnGlnArgValArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||
Db 613 CAGGCGGAGGCGGATGAGGAGCGCGGCTCTCAGCGCCATCCGCGAGCCCTGGGG 672
QY 201 ProLeuValGluGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
|||||
Db 673 CCCCTGTGGAACAGGGCGCGCTGCGGGCGCCACTGTGGGCTCCCTGGCGGCGAGCCCT 732
QY 221 LeuGlnGluArgAlaGlnAlaThrGlyGluArgLeuArgAlaArgMetGluGln-MetG 240
|||||
Db 733 A--CAGAGACGCGGCGCAGGCTCGGCGAGCGGTGGCGCGCGATGAGAGAAATGGG 790
QY 240 ySerArgThrArgAspArgLeuAspGluValLysGlnGlnAlaGlnValArgAlaLys 260
|||||
Db 791 CAGCCGAGCCGCGGAGCG--CTGGAGCAAGTGAAGACACAGTGGCGGAGAGG-TGGCCCAA 847


```

QY 260 sLeuGluc1a1aGln 266
|||||:|||||:
Db 848 GCTGAGACGACGACGAG 866

RESULT 14
BM042228
LOCUS BM042228 757 bp mRNA linear EST 07-NOV-2001
DEFINITION 603616186F1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5557004 5',
mRNA sequence.
ACCESSION BM042228
VERSION BM042228.1 GI:16771495
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 757)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2021 row: k column: 21
High quality sequence stop: 757.
FEATURES
source
location/Qualifiers
1..757
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5557004"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOT87; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."
BASE COUNT 134 a 229 c 236 g 98 t
ORIGIN
Alignment Scores:
Pred. No.: 1,16e-108 Length: 757
Score: 1131.00 Matches: 234
Percent Similarity: 98.73% Conservative: 0
Best Local Similarity: 98.73% Mismatches: 2
Query Match: 71.18% Indels: 2
Gaps: 0
US-09-827-854-18 (1-317) x BM042228 (1-757)
QY 1 MetLysValLeuTrpAlaAlaLeuValThpPheLeuAlaGlyCysGlnAlaLysVal 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 50 ATGAAGTTCCTGTCGGCTGCTGTCATTCCTGTCAGAGATGCCAGGCCAAGGTG 109
QY 21 GluGlnAlaValGluTrpGluGluProGluProGluGluGlnGlnTrpGlnGlnSer 40
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 GAGCAAGCGGTGGAGAGAGAGCCGAGCCGAGCTGGCCAGACAGCCGAGTGCAGAGC 169
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 170 GGCACGCGCTGGAGACAGCTGGCTTTGGGATTACTGCGCTGGGTGCAGACA 229

```

```

QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 230 CTGCTCAGAGAGGTGAGAGAGAGACTCTCAGCTCCAGGTCCAGGAGCAAGAGGCG 289
QY 81 LeuMetAspGluTrpMetLysGluLeuValAlaTyrLysSerGluLeuGlnGluGlnLeu 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 290 CTGATGACAGACAGACCATGAAGGACTTGAAGCCCTACAAATCGGAATCGAGAGCAACTG 349
QY 101 ThrProValAlaGluGluTrpArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 350 ACCCGGTGGGAGAGAGAGCGGGGACAGGCTGTCCAAAGAGAGTCCAGCGCGGACAGCC 409
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 410 CGGCTGGCGCGGACAGAGAGAGCTGTGGCGCGCTGTGACGTACCCCGGAGAGTG 469
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 470 CAGCCATGCTCGGCGCAGACACCGAGGAGCTGGGGTGGCTCCCTCCACCTGCGC 529
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 530 AAGCTGGTAAGGCGGCTCTCCGGAGTCCGATGACCTGCAGAAAGCGCTGGCAGGTGAC 589
QY 181 GlnAlaGlnAlaArgGlnGlnGlnAlaGlnArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 590 CAGCCCGGGGCGCGGAGGCGCGGAGCGCGCTCAGCCCATCCCGCGAGCGCTGGGG 649
QY 201 ProLeuValGlnGlnArgVal-ArgAlaAlaThrValGlySerLeuAlaGlnPr 220
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 650 CCCCTGGTGGAGACAGGCGCGCTGGCGGCGCCACCTGTGGCTCCTGGCCGCGCAGCC 709
QY 220 OLeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMet 236
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 710 GCTACAGAGAGCGGCG-CAGGCTGGGGGAGAGGCTCCGCGCGGATG 757

RESULT 15
LOCUS BG706129 907 bp mRNA linear EST 07-MAY-2001
DEFINITION 602669093F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4792030 5',
mRNA sequence.
ACCESSION BG706129
VERSION BG706129.1 GI:13981169
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 907)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10669 row: 1 column: 23
High quality sequence stop: 832.
FEATURES
source
location/Qualifiers
1..907
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4792030"
/tissue_type="hypothalamus"
/note="Organ: Brain; Vector: pBluescriptR (modified)

```


GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:08:57 ; Search time 2307.23 Seconds
(without alignments)
3998.551 Million cell updates/sec

Title: US-09-827-854-19
Perfect score: 1585
Sequence: 1 MKVLMAALLVFLACQAKV.....VEKVOAAVGTSAAPVSDNH 317

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=rlp
-Q=/cgn2.1/USPRO.spool/US09827854/rlmat.11032003.101610.27486/app.query.fasta.1.3576
-DB=genbml -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human4.cdi -LIST=45
-DOCALIGN=200 -THR.SCORE=pcr -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEADSIZE=5001 -MINLEN=0 -MAXLEN=2000000000
-USER=US09827854.ecgn.1.1.13008.arunat.11032003.101610.27486 -NCPU=6 -ICPU=3
-NO_XIPXY -NO_MMAP -LARGEJUNKER -NEGSCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: genbml:*
2: gb_ba:*
3: gb_hlg:*
4: gb_in:*
5: gb_om:*
6: gb_ov:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_in:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pin:*
35: em_hlg_rod:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_hlgo_hum:*
40: em_hlgo_mus:*
41: em_hlgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1581	99.7	1110	6	E00359	E00359 cDNA coding
2	1581	99.7	1110	6	E00823	E00823 DNA sequence
3	1581	99.7	1147	6	AX302545	AX302545 Sequence
4	1581	99.7	1156	6	BD004278	BD004278 Apo E hum
5	1581	99.7	1156	9	HUMAP0E3	K00396 Homo sapien
6	1581	99.7	1186	9	BC003557	BC003557 Homo sapi
7	1574	99.3	1110	6	E08423	E08423 DNA coding
8	1573	99.2	1156	6	BD004277	BD004277 Apo E hum
9	1572	99.2	1157	6	AX333278	AX333278 Sequence
10	1572	99.2	1157	6	AX409597	AX409597 Sequence
11	1572	99.2	1157	6	I15975	I15975 Sequence 1
12	1572	99.2	1157	9	HUMAP0E	M12529 Human apoli
13	1569	99.0	1156	6	BD004279	BD004279 Apo E hum
14	1447	91.3	1178	9	MFAPOE	X13887 Monkey mRNA
15	1392.5	87.9	5491	9	AF261279	AF261279 Homo sapi
16	1392.5	87.9	5491	6	AX358722	AX358722 Sequence
17	1392.5	87.9	41907	9	AF050154	AF050154 Homo sapi
18	1392.5	87.9	107567	9	AC011481	AC011481 Homo sapi
19	1375	86.8	5515	9	HUMAP0E4	M10065 Human apoli
20	1366	84.9	5413	9	AF261280	AF261280 Pan trogl
21	1263	79.7	4762	9	BABAP0E	M29322 Baboon apoli
22	1262	79.6	208239	2	AC021988	AC021988 Homo sapi
23	1254	79.1	1138	4	AF303830	AF303830 Tupai a gl
24	1166	73.6	1060	4	RABAP0LP	M36603 Rabbit apoli
25	1159.5	73.2	965	6	AX384545	AX384545 Sequence
26	1159.5	73.2	1108	4	BTAPOLP	X61171 B. taurus mr
27	1159.5	73.2	5617	6	AX384541	AX384541 Sequence
28	1159.5	73.2	6026	6	AX384539	AX384539 Sequence
29	1147	72.4	718	9	AF20049753	AF200497 Pan trogl
30	1145	72.2	1045	10	MUSAP0E	M12414 Mouse apoli
31	1145	72.2	1104	10	BC028816	BC028816 Mus muscu
32	1143	72.1	718	9	AF20050053	AF200500 Gorilla g
33	1143	72.1	718	9	AF20050353	AF200503 Pongo pyg
34	1141.5	72.0	1154	9	BTAPOEMR	AF200505 B. taurus mr
35	1139	71.9	718	9	AF20050653	AF200506 Hylobates
36	1139	71.9	1122	4	SSAP0E	X72835 S. scrofa mr
37	1134	71.5	1126	6	ARI64342	ARI64342 Sequence
38	1134	71.5	1126	6	AR205885	AR205885 Sequence
39	1090	68.8	951	10	S76779	S76779 RAP0E-apoli
40	1068	67.4	959	10	MUSAP0E	M73490 Mus musculu
41	1052.5	66.4	1069	10	RATAP0E	J00705 rat apoli
42	1025	64.7	228698	2	AC127479	AC127479 Mus muscu
43	1025	64.7	237653	2	AC073760	AC073760 Mus muscu
44	999	63.0	4856	10	MUSAP0E	D00466 Mus musculu
45	998	63.0	4267	6	ARI64387	ARI64387 Sequence

RESULT 1

ALIGNMENTS

E00359
 LOCUS E00359 1110 bp RNA Linear PAT 29-SEP-1997
 DEFINITION cDNA coding human apolipoprotein E3.
 ACCESSION E00359
 VERSION E00359.1 GI:2168646
 KEYWORDS JP 1985118189-A/1.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1110)
 Teranishi, Y., Takamatsu, N., Matsui, Y., Kimura, M. and Ikeda, Y.
 REFERENCE
 AUTHORS DNA FRAGMENT
 TITLE Patent: JP 1985118189-A 1 25-JUN-1985;
 JOURNAL MITSUBISHI CHEM IND LTD
 COMMENT OS human
 PN JP 1985118189-A/1
 PD 25-JUN-1985
 PF 29-NOV-1985 JP 1983224980
 PI TERANISHI YUTAKA, TAKAMATSU NOBUHIKO, MATSUI YASUSHI, PI
 KIMURA MASAKO,
 IC Ikeda Yasuko
 PC C12N15/00,C07H21/04//C12P21/00;
 CC strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC *source: tissue_type=liver;
 FH key Location/Qualifiers
 FT CDS 15..968
 FT sig_peptide 15..68
 FT mat_peptide 69..965
 FT /product='apolipoprotein E3 precursor' FT
 FT /product='apolipoprotein E3 signal peptide' FT
 FT Location/Qualifiers
 1..1110
 /organism='Homo sapiens'
 /db_xref='taxon:9606'
 BASE COUNT 198 a 353 c 416 g 143 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.65e-80 Length: 1110
 Score: 1581.00 Matches: 316
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.68% Mismatches: 0
 Query Match: 99.75% Indels: 0
 DB: 6 Gaps: 0
 US-09-827-854-19 (1-317) x E00359 (1-1110)
 QY 1 MetlysvallleutrpAlaAlaLeuLeuValThrLeuAlaGlyCysGlnAlaLysVal 20
 Db 15 ATGAAGGTTCCTGGTGGCTGCTGCTGCACATTCCTGGCAGATGCCAGGCCAAGTG 74
 QY 21 GluAlaAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 Db 75 GAGCAACCGGTGGAGACAGACCCGAGCTGGCGCCACGACGACGAGCTGGCAAGC 134
 QY 41 GlyLlnArGTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 Db 135 GGCAGCGCTGGGAACTGGCAGCTGGTGGCTTTGGGATTACCTGGCTGGGTGCACACA 194
 QY 61 LeuSerLluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 Db 195 CTGCTGTGAGAGGTGCGAGGAGGAGCTCTCAGCTCCAGGTCACCCAGGAAGTCAAGGCG 254
 QY 81 LeuLeuAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100
 Db 255 CTGATGACAGACGACCATGAAGAGGTTGAAGGCCCTACAAATCGACACTGGAGGAACAACCTG 314

QY 101 ThrProValAlaGluGluThrThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
 Db 315 ACCCGGCGCGGAGAGACCGCGGCGACGGCTGTCCAAAGGAGTGCAGAGCGCGCGACAGG 374
 QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 Db 375 CGAGCTGGCGCGGACATGAGAGACCTGCGCGCGCTGGTGCATGACCCGCGGAGAGTG 434
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 435 CAGGCCATGCTGGGCCAGAGACGACGAGAGAGTGGGGTGGCCCTCGGCTCCACACTGCC 494
 QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 Db 495 AAGCTGGTAAAGCGGCTCTCCGCGATGCCGATGACCTGCAGAAAGCCCTGGCAGTAC 554
 QY 181 GlnAlaGlyAlaArgGluGluGluAlaGlnArgGlyLeuSerHisLeuArgGluGly 200
 Db 555 CAGCGCGGGCGCGCGAGGGCGCGAGCGCGGCTCCAGCGCATCCCGAGCGCGCTGGGG 614
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 615 CCCCTGGTGAACAGGCGCGCGTGGCGCGCGCTGGGCTCCCTGGCGCGCGAGCG 674
 QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
 Db 675 CTACAGAGAGCGGCGCGCGCTGGCGGAGCGCGCTGGCGCGCGGATGAGAGATGGCG 724
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAlaArgLys 260
 Db 735 AGCGGAGCGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 794
 QY 261 LeuGluGluGlnAlaGlnGlnLysArgLeuGlnAlaGluValAlaPheGlnAlaArgLeuLys 280
 Db 795 CTGAGAGAGCAGCGCGCGAGATGCGCTGCGAGCGCGAGCGCTTCCAGCGCGCGCTAAG 854
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 Db 855 AGCTGTTTCAGAGCGCGCTGTGGAAGACATGACAGCGCGCGCGCGCTGTGTGAGAGAG 914
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 915 GTGAGGCTCGCGCGGACACGCGCGCTGTGTGCCAGCGCAATCAC 965
 RESULT 2
 E00823 1110 bp DNA Linear PAT 29-SEP-1997
 LOCUS E00823
 DEFINITION DNA sequence coding for human apolipoprotein E and its signal peptide.
 ACCESSION E00823
 VERSION E00823.1 GI:2169084
 KEYWORDS JP 1986096997-A/1.
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1110)
 AUTHORS Teranishi, Y., Matsui, Y., Ikeda, Y. and Kimura, M.
 TITLE PRODUCTION OF HUMAN APOLIPOPROTEIN E-LIKE PROTEIN
 JOURNAL Patent: JP 1986096997-A 1 15-MAY-1986;
 MITSUBISHI CHEM IND LTD
 COMMENT OS Human (Homo sapiens)
 PN JP 1986096997-A/1
 PD 15-MAY-1986
 PF 16-OCT-1984 JP 1984216987
 PI TERANISHI YUTAKA, MATSUI YASUSHI, IKEDA YASUKO, KIMURA MASAKO
 PC C12P21/00,A61K37/74,A61K37/04,C12N15/00,(C12P21/00,C12R1:19),
 PC (C12N15/00,
 PC C12R1:19);
 CC strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC *source: tissue_type=liver;

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------

QY 121 ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
Db 406 CGCGTGGGCGCGGACATGAGAGAGCTGTGGCGCGCTGGTGGTGCAGTACGGCGGAGGTG 465
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 466 GAGCCATGCTCGCCCAAGACAGCAGAGAGCTGGGGTGGCGCTCGCTCCACCTGGCG 525
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnTyrArgLeuAlaValTyr 180
|||||
Db 526 AAGCTGCTTAAGCGGCTCTCTCCGATGCCATGACCTGCAGAACCGCTTGGCACTGTAC 585
QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||
Db 586 CAGCGCGGCGCGCGGAGCGCGCGAGCGCGCTCCACGCGCTCCGCGAGCGCTTGGCG 645
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
|||||
Db 646 CCCCTGGTGGAAACAGGGCGCGGTGGGGCGCCACTGTGGCTCTCCGCGCGCGAGCGG 705
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
|||||
Db 706 CTACAGAGAGCGGGGCCAGGCTGGGGCGAGCGGCTGGCGCGCGATGGAGAGATGGGC 765
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
|||||
Db 766 AGCGGAGCCCGCGAGCGCGCTGGAGAGGTGAGGAGCGAGGTGGCGGAGGTGGCGG 825
QY 261 LeuGluGlnGluAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
|||||
Db 826 CTGGAGGAGAGCGCCGACAGATAGCTGAGCGCGAGCGCTTCCAGCGCGCGCTTCAAG 885
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
|||||
Db 886 AGCGGTTCGAGCGCCCTGTGTGAAGACATGCAGCGCCGATGGCGGCGCTGTGTGAAG 945
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
|||||
Db 946 GTGCAGCGTCCGCGTGGGAGCAGCAGCGCCCGCTGTGCCAGCAGCAATCAC 996
RESULT 4
BD004278 1156 bp DNA linear PAT 31-JAN-2002
LOCUS BD004278 Apo E humanized mammal.
DEFINITION BD004278
ACCESSION BD004278.1 GI:18632239
VERSION BD004278.1 GI:18632239
KEYWORDS JP 2001017028-A/2.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1156)
AUTHORS Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.
TITLE Apo E humanized mammal.
JOURNAL Patent: JP 2001017028-A 2 23-JAN-2001;
MITSUBISHI CHEMICAL CORP
OS Homo sapiens (human)
PN JP 2001017028-A/2
PD 23-JAN-2001
PR 28-APR-2000 JP 2000128919
COMMENT
PI SHINOBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA PC
A01K67/027, A61K45/00, A61P25/28, A61P43/00, C12N5/10, PC
C12N5/09//C07K14/775,
PC (C12N5/10, C12R1:91), C12N5/00, C12N15/00, (C12N5/00, C12R1:91) CC
FEATURES
source FH Key Location/Qualifiers
FT CDS (61)..(1011).
1..1156 Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 208 a 368 c 432 g 148 t
ORIGIN
Alignment Scores:
Pred. No.: 5,89e-80 Length: 1156
Score: 1581.00 Matches: 316
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.68% Mismatches: 0
Query Match: 99.75% Indels: 0
DB: 6 Gaps: 0
US-09-827-854-19 (1-317) x BD004278 (1-1156)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
Db 61 ATGAAGGTTCTGTGGGCTGCGCTGCTGTCATCTCTGCAGAGATGCAGGCCAAGGTG 120
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
|||||
Db 121 GAGCAAGCGGTGGAGACAGACCGGAGCCGAGCTGCCCGCAGACAGAGTGCAGAGC 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
Db 181 GGCAGCGCTGGCAACTGGCACTGGGTGCTTTTGGATTAACCTGCGCTGGGTGCAGACA 240
QY 61 LeuSerGluGlnValGlnGluGluLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
Db 241 CTGTCAGAGAGGTGACAGAGAGAGCTGCTCAGTCCAGGTCCACCCAGAGAACTAGGGCG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100
|||||
Db 301 CTGATGAGCAGAGACCATTAAGAGCTGAAGGCTTCAAAATCGAATCGAAGAACTAG 360
QY 101 ThrProValAlaGluGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
|||||
Db 361 ACCCGGTGGGAGAGAGACACCGCGGACAGCGCTGTCCAAAGAGACTGCAGCGCGGAGGCC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
Db 421 CGCGTGGGCGCGGACATGAGAGAGGTGGCGCGCGCTGTGGTGCAGTACCGCGGAGGTG 480
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 481 CAGCGCATGCTGGCCCAAGACAGCAGAGAGCTGGCGGCGCTCCGCTCCACCTGGCG 540
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnTyrArgLeuAlaValTyr 180
|||||
Db 541 AAGCTGCTTAAGCGGCTCTCTCCGATGCCATGACCTGCAGAAAGCGCTGGCACTGTAC 600
QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||
Db 601 CAGCGCGGCGCGCGGAGCGCGGAGCGCGCTCCACGCGCTCCGAGCGCGCTGGGG 660
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
|||||
Db 661 CCCCTGTGGAAACAGGGCGCGCTGGCGCGCGCCACTGTGGCTCTCCGCGCGGAGCGG 720
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
|||||
Db 721 CTACAGAGAGCGGGCCAGGCTGGGGGAGCGGTGGCGCGCGGATGGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
|||||
Db 781 AGCGGAGCCCGCGAGCGCGCTGGAGGTGAAGAGCAGGTGGGGAGGTGGCGGCAAG 840
QY 261 LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
|||||
Db 841 CTGGAGAGAGCGGCCAGGCTGGGGGAGCGGTGGCGCGCGGATGGAGAGATGGGC 900
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
|||||
Db 901 AGCTGTTTCGAGCGCCCTGTGTGAAGACATGCAGCGCCCAAGTGGCGGCTGTGTGAAG 960

QY	301	ValGlnAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db	961	GTGCAGGCTGCGTGGGACACAGCGCCGCTGTGTCAGCAGCAATCATC 1011
RESULT 5		
HUMAPOE3		
LOCUS		
DEFINITION	Homo sapiens preapolipoprotein E (APOE) mRNA, complete cds.	1156 bp mRNA linear PRI 24-NOV-2000
ACCESSION	K00396	
VERSION	K00396.1 GI:178850	
KEYWORDS	apolipoprotein; apolipoprotein E; lipoprotein; polymorphism; very low density lipoprotein.	
SOURCE	Homo sapiens.	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE		
AUTHORS	1 (bases 355 to 1156) Breslow,J.L., McPherson,J., Nussbaum,A.L., Williams,H.W., Loftquist-Kahl,F., Karathanasis,S.K. and Zannis,V.I. Identification and DNA sequence of a human apolipoprotein E cDNA clone	
TITLE		
JOURNAL	J. Biol. Chem. 257 (24), 14639-14641 (1982)	
MEDLINE	83082756	
PUBMED	6897404	
REFERENCE		
AUTHORS	2 (bases 250 to 777) Wallis,S.C., Rogne,S., Gill,L., Markham,A., Edge,M., Woods,D., Williamson,R. and Humphries,S. The isolation of cDNA clones for human apolipoprotein E and the detection of apoE RNA in hepatic and extra-hepatic tissues	
TITLE		
JOURNAL	EMBO J. 2 (12), 2369-2373 (1983)	
MEDLINE	84131952	
PUBMED	6199196	
REFERENCE		
AUTHORS	3 (bases 1 to 1156) Zannis,V.I., McPherson,J., Goldberger,G., Karathanasis,S.K. and Breslow,J.L. Synthesis, intracellular processing, and signal peptide of human apolipoprotein E	
TITLE		
JOURNAL	J. Biol. Chem. 259 (9), 5495-5499 (1984)	
MEDLINE	84185684	
PUBMED	6325438	
REFERENCE		
AUTHORS	4 (bases 88 to 1156) McLean,J.W., Elshourbagy,N.A., Chang,D.J., Mahley,R.W. and Taylor,J.M. Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing of a new variant	
TITLE		
JOURNAL	J. Biol. Chem. 259 (10), 6498-6504 (1984)	
MEDLINE	84212473	
PUBMED	6327682	
REFERENCE		
AUTHORS	5 (bases 577 to 624) Gill,L.L., Peoples,O.P., Pearson,D.H., Robertson,F.W., Humphries,S.E., Cumming,A.M. and Hardman,N. Isolation and characterisation of a variant allele of the gene for human apolipoprotein E	
TITLE		
JOURNAL	Biochem. Biophys. Res. Commun. 130 (3), 1261-1266 (1985)	
MEDLINE	85279526	
PUBMED	2992507	
REFERENCE		
AUTHORS	6 (sites) Rall,S.C. Jr., Newhouse,Y.M., Clarke,H.R., Weisgraber,K.H., McCarthy,B.J., Mahley,R.W. and Bersot,T.P. Type III hyperlipoproteinemia associated with apolipoprotein E phenotype E3/3. Structure and genetics of an apolipoprotein E3 variant	
TITLE		
JOURNAL	J. Clin. Invest. 83 (4), 1095-1101 (1989)	
MEDLINE	89198059	
PUBMED	2539388	
COMMENT		
	[1] corrected in [J. Biol. Chem. 258, 11422-11422 (1983)]. [J. Biol. Chem. 258, 11422-11422 (1983)] correction of [1]. [4] epsilon-3 and variant. [5] epsilon-2 allele. [6] sites; mutations resulting in type III hyperlipoproteinemia. Apo E is a component of normal human very low density lipoprotein. There are six human apo E phenotypes known to result from a single structural gene, three of the common alleles being epsilon-4,	

FEATURES	
source	epsilon-3 and epsilon-2. This sequence appears to be of the epsilon-3 allele. [1] argues that the apo E polymorphism involves mutations in the structural coding region; for example the epsilon-2 phenotype which is characterized by hyperlipoproteinemia is thought to result from a c to t change (arg to cys) at base 586 below [3],[5]. The sequence shown is 57% homologous with human apo A-I and 81% homologous with rat apo E. For the epsilon-4 sequence, see the separate entry. [J. Biol. Chem. 258, 11422-11422 (1983)] is too extensively revised relative to [1] to record the revisions in the FEATURES table. The sequence below is that of [J. Biol. Chem. 258, 11422-11422 (1983)] [3]. Apo E is located on chromosome 19 -- Jackson, Bruns and Breslow, PNAS USA 81, 2945-2949 (1984) -- and is believed to be linked to the Apo C-II gene (see separate entry). The two mutations causing type III hyperlipoproteinemia (apolipoprotein E phenotype E3/3) produces substitutions of Arg for Cys at amino acid 112 and Cys for Arg at amino acid position 143. Draft entry and printed copy of sequence for [1] kindly provided by L.L. Gill, University of Aberdeen. Complete source information: Human liver [1],[J. Biol. Chem. 258, 11422-11422 (1983)]. [2], [3], [4] and blood [5], cDNA to mRNA. Location/Qualifiers 1..1156 /organism="Homo sapiens" /db_xref="taxon:9606" /map="19q13.2" /tissue_type="liver and blood" 1..1156 /gene="APOE" 61..1014 /gene="APOE" /codon_start=1 /product="preapolipoprotein E" /protein_id="AAB59546.1" /db_xref="GI:178851" /db_xref="GDB:600-119-691" /translation="MKVLAALLVTFLLAGCQAKVQAVETPEPELRQGTMOGQRMVALGRFMDYLRWVOTLSQVDELLSQTVELRLAMDETKEKAKSLEEDLTPVAEETRARLSKWTQAQAQARLGDMDVCGRLVQYGEVQAMGOSTEELRVLASHLRKRRLRDADDLQRLRAYQAGAREGASLRERLGPVGGQRYVAATVGSIAQOPLQERQAWMGRLRRARMEGMSRTRDIDEYKEQVAERAKLEEQAOQIRIQANAFAARKSWPEPLVYEDMQRQNAGLYEVQAAGVTSAAFPVSDNH" 61..114 /gene="APOE" /note="apolipoprotein E" 115..1011 /gene="APOE" /product="apolipoprotein E mature peptide" 267 /gene="APOE" /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)]. [3], [4]" /citation=[2] /replace="c" 279 /gene="APOE" /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)]. [3]" /citation=[4] /replace="a" 288 /gene="APOE" /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)]. [3]" /citation=[4] /replace="a" 369 /gene="APOE" /note="in variant" /citation=[4] /replace="a" 409
gene	
CDS	
sig_peptide	
mat_peptide	
variation	
variation	
variation	
variation	
variation	


```

/ gene="APOE"
/ note="in variant"
/ citation=[4]
/ replace="a"
448
variation
/ gene="APOE"
/ note="in type III hyperlipoproteinemia"
/ replace="c"
538
variation
/ gene="APOE"
/ note="in type III hyperlipoproteinemia"
/ replace="t"
568
variation
/ gene="APOE"
/ note="in variant"
/ citation=[4]
/ replace="c"
586
variation
/ gene="APOE"
/ note="in epsilon-2 allele"
/ citation=[3]
/ citation=[5]
/ replace="t"
702
variation
/ gene="APOE"
/ note="c in [J. Biol. Chem. 258, 11422-11422 (1983)]. [3], [4]"
/ citation=[2]
/ replace="g"
783
variation
/ gene="APOE"
/ note="in variant"
/ citation=[4]
/ replace="t"
858
variation
/ gene="APOE"
/ note="g in [J. Biol. Chem. 258, 11422-11422 (1983)]. [3] and [4] variant; a in [4] normal"
/ citation=[4]
/ replace="a"
368 c 432 g 148 t

BASE COUNT 208 a 368 c 432 g 148 t
ORIGIN

Alignment Scores:
Pred. No.: 5.89e-80 Length: 1156
Score: 1581.00 Matches: 316
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.68% Mismatches: 0
Query Match: 99.75% Indels: 0
DB: 9 Gaps: 0

US-09-827-854-19 (1-317) x HUMAPOB3 (1-1156)
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 61 ATGAAGGTTCTGTGGTGGCTGCTGTCACATTCCTGGCAGATGCCAGGCCAAGTGT 120
QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 121 GAGCAACGGGTGGAGACAGACCCGAGCTGCGCCACGACGACCGAGTGGCAGAGC 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 181 GGCACGCGCTGGACACTGGCTGCTTTGGGATTAACCTGCCCTGGGTGCAGACA 240
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 241 CTGTCTGAGAGGTGCGAGGAGAGCTGCTCAGCTCCAGGTACACCGAGAACTGAGGGCG 300
QY 81 LeuMetAspGluTrpMetLysGluLeuValAlaTyrLysSerGluLeuGluGlnGlnLeu 100
DB 301 CTGATGACAGACGACATGAAGGAGTGAAGGCCCTACAAATCGAATGAGAGCAACTG 360

```

```

QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 361 ACCCGGTGGCGGAGAGACCGGGGACAGCGCTGTCCAAAGAGCTGCAGAGCGCGCAGGCC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 421 CGGCTGGGCGGACATGAGAGACGTGTGCGCCCTGGGTGCAATGCCGGGGAGAGTG 480
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGCCATGCTCGGCCACAGACGACGAGAGAGTGGGGTGGCCCTCGCTCCACCTCGC 540
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspLeuGlnLysArgLeuAlaValTyr 180
DB 541 AAGCTGGTAAAGCGCTCTCCGCGATGCCATGCGAGAGCCCTGGCAGTGTAC 600
QY 181 GlnAlaGlyAlaArgGluGluGluArgGlyLeuSerAlaLeuArgGluLeuGly 200
DB 601 CAGCGCGGGCGCGGAGGGCGGAGCGGCTCAGCGCATCCCGAGCGCGCTGGG 660
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 661 CCCCTGGTGGACAGGGCCCGGTGCGGCGCCACCTGTGGGTCCCTGGCGCGGACGCG 720
QY 221 LeuGlnGluArgAlaGlnAlaTyrGlyGluArgLeuArgAlaArgMetGluMetGly 240
DB 721 CTACAGAGAGCGGGCCAGCGCTGGGCGAGCGGCTGCGCGGAGTGAAGATGGCG 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 781 AGCCGAGCCCGGACCCCTGGAGAGGTGAAGAGCAGGGGGGAGGTGGCGGCCAAG 840
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 841 CTGGAGAGGACGAGGCCAGCATATGCTGACAGCCAGCGCTCCAGCGCGGCTCAAG 900
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 901 AGCTGTTCAGAGCCCGCTGGGTGAAGACATGACAGCGCCAGTGGCGGTGTGGAGAG 960
QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspHis 317
DB 961 GTGCAGGCTGCGGTGGACACGCGCCCTGTGTGCCAGGCACATTCAC 1011

RESULT 6 1186 bp mRNA linear PRI 12-JUL-2001
BC003557 BC003557
LOCUS Homo sapiens, apolipoprotein E, clone MGC:1571 IMAGE:3355712, mRNA,
DEFINITION
ACCESSION BC003557.1 GI:13097698
VERSION BC003557
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1186)
REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT Contact: MGC help desk
Email: c9apds-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
Info@bcsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,

```


Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>
Series: IRM1 Plate: 6 Row: h Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA q1: 4557324.

1. .1186

CDS

248 a 366 c 425 g 147 t

248 a 366 c 425 g 147 t

Query Match:

US-09-827-85

QY	1	MeysValleuTrpAlaAlaIleuValThrPheIeuNlaGlyGlySlnAlaIleVal	20
Db	46	ATGAAGTTCGTGGCGTCGCTCTGTGCATCTCTGGAGATCCAGGCCAAGGTG	105
QY	21	GIuGlnAlaValGIuThrGIuProGlnProGIuIeuArgSlnGlnThrGIuTrpGlnSer	40
Db	106	GAGCAAGCGGTGGAGACAGACCAGCCAGCTGGCCAGCAGACCGAGTGGCAAGC	165
QY	41	GIyGlnArgTrpGluIeuAlaIleuGlyIArgPheTrpAspIlyIleuArgTrpValGlnThr	60
Db	166	GGCCAGCGCTGGGAACCTGGCACTGGGCGCTTTGGGATTACCTGGCCCTGGGTCCACACA	225
QY	61	LeuSerGIuGlnAlaGlnGIuIleuLeuSerSerGIuAlaThrGlnGlnIleuArgAla	80
Db	226	CTGTCTAGCAGAGTGCAGAGAGAGACTCTCAAGTCCCAAGTCAACCCAGAACTAGAGGCG	285
QY	81	LeuMetAspGIuThrMetIySGluIeuIySAlaTrpIySerGIuIeuGlnGIuGlnLeu	100
Db	286	CTGATGGACGACCATCAAGAGAGATTGAGCGCTACAAATCGCAACTGGAGAACCACTG	345
QY	101	ThrProValAlaGIuGIuThrArgAlaArgIeuSerIySGluIeuGlnAlaIaGlnAla	120
Db	346	ACCCCGGTGGGGAGAGACCGCGGGCAAGGCTGTCCAAAGGAGCTGCAGCGGGCGCCAGGCC	405
QY	121	ArgIeuGlnIyAlaAspMetGIuAspValIySGlyIArgIeuValGlnTrpArgGIuGlnVal	140

Dh	406	CGGCTGGGCGGACATGGAGACGCTGGCGCGCTCGTGGACGTACCGGCGGAGGTG	465
Qy	141	GLNALametiLeuglYglInSerThrGluGluLeuArGValArGLeuAlaSerHisLeuArg	160
Dh	466	CAGGCGATGCTGGCCAGACAGCACCGAGAGAGTGGGGGCGCCCTCCCTCCACCTTGGCG	525
Qy	161	LysLeuArGlnArGLeuLeuArGAspAlaAspAspleuGlnLysArgLeuAlaValTyr	180
Dh	526	AAGTGGCGTAAAGCGGCTCTCCGGATGCCATGACCTGCAGAAAGCGCTGGCACTGTAC	585
Qy	181	GLNALaGlyAlaArGluGluYAlaGluArgLysLeuSerAlaIleArgGluArgLeuGly	200
Dh	586	CAGCGCGGGGCGCCGAGAGCGCGAGGCGCGCTCAGCGCCATCCGCGACGCGCTGGGG	645
Qy	201	ProLeuValGluGluGlnArgValArgAlaAlaIleThrValGlySerLeuAlaGluGlnPro	220
Dh	646	CCCTGTGGTGAACAGGGCGCGCTGGGGCGCCCGACCTGTGGGCTCCCTGGCCGGCAGCCG	705
Qy	221	LeuGluGluArGAlaGlnAlaIleTrpGluGluArgLeuArGAlaArgMetGluGluMetGly	240
Dh	706	CTACAGGAGCGGGGCCAGAGGCTCTGGGGCGAGGGCGCTCGCGCGGATGAGAGATGGGC	765
Qy	241	SerArgThrArgAspArgLeuAspGluValLysGluGluGlnValAlaGluValArgAlaLys	260
Dh	766	AGCGGAGCCCGGACCGGCTGTGAGAGGTGAAGGAGCGAGTGGCGGAGTGGCGCCCAAG	825
Qy	261	LeuGluGluGlnAlaGlnGlnIleArgArgLeuGlnAlaGluAlaIlePheGlnAlaArgLeuLys	280
Dh	826	CTGGAGGAGCGAGGCCAGACAGATACGCCCTTCGAGGGCGAGGCGCTTCACAGCCCGCTCAAG	885
Qy	281	SerTrpPheGluProLeuValGluAspMetGlnArgGlnIleTrpAlaGlyLeuValGluLys	300
Dh	886	AGCGGTTCGAGCCCTGGTGTGAAGACATGACGCCCGTGGGCGGGCTGTGGAGAAG	945
Qy	301	ValGlnAlaAlaValGlyThrSerAlaAlaIleProValProSerAspAsnHis	317
Dh	946	GTGGAGGCTGGCTGGGGGACACAGGCGCGCTGTGGCTGGCCAGCAACATCAC	996

LOCUS	1110 bp	RNA	linear	PAT 29-SEP-1997
DEFINITION	E08423 DNA coding human apolipoprotein E3.			
ACCESSION	E08423			
VERSION	E08423.1 GI:2176540			
KEYWORDS	JP 1994315392-A/1.			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
AUTHORS	1 (bases 1 to 1110)			
TITLE	Morimoto,H. and Teranishi,Y.			
JOURNAL	METHOD FOR PRODUCING APOLIPOPROTEIN Patent: JP 1994315392-A 1 15-NOV-1994; MITSUBISHI KASEI CORP			
COMMENT	OS Homo sapiens (human)			
	PN JP 1994315392-A/1			
	PD 15-NOV-1994			
	PF 11-JUN-1985 JP 1994015433			
	PI MORIMOTO HIRONORI, TERANISHI YUTAKA			
	PC C12P21/02.C07K13/00.C12N5/10.C12N5/12.C12N5/18.(C12P21/02, C12P1.91).			
	PC (C12N5/10,C12P1.91);			
	CC strandedness: Double;			
	CC topology: Linear;			
	Key Location/Qualifiers			
FT	1..1110			
FT	source			
FT	/organism='Homo sapiens'			
FT	sig_peptide 15..68			
FT	mat_peptide 69..968			
FT	/product='human apolipoprotein E3'.			
FEATURES	Location/Qualifiers			
source	1..1110			

Db 361 ACCCGGTGGCGGAGACGCCGCGCACGGCTGTCCAAAGNAGCTGACGGCGCGCAGGCC 420
 Oy 121 ArgLeuGluValAlaAspMetGluAspValCysGlyArgLeuValGlnGlyArgGlyVal 140
 Db 421 CGGCTGGCGCGGACATAGGAGAGAGTGTGGCGCCCTGGGTGAGTACCGCGGAGGTG 480
 Oy 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 481 CAGGCATGCTCGGCGCAGAGACGAGAGAGTGGGGTGGCGCTCGGCTCCACCTCGCC 540
 Oy 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspLeuGlnLysArgLeuAlaValTyr 180
 Db 541 AAGCTGGTAAAGCGGCTCTCTCCGCGATGCGATGACCTGCAAGAGTCCCTGGCAGTAC 600
 Oy 181 GlnAlaGlyAlaArgGlnGluGlnAlaGlnGlyLeuSerAlaIleArgGlnArgLeuGly 200
 Db 601 CAGCGCGGGGCGCGCGAGGCGCGGCGCTCAGCGCCATCCGCGAGCGGCTGGGG 660
 Oy 201 ProLeuValGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
 Db 661 CCCCTGGTGMAAGAGGCGCGGCTGGGCGCGCAGCTGGGGCTCCCGCGGCGCAGCGG 720
 Oy 221 LeuGlnGluArgAlaGlnAlaATTPGlyGluArgLeuArgAlaArgMetGlnGlnMetGly 240
 Db 721 CTACAGAGAGCGGGCCAGCGCTGGGGCGAGCGGCTGCGCGCGCGGATGAGAGATGGGC 780
 Oy 241 SerArgThrArgAspArgLeuAspGluValGlyGlnGlnAlaGlnValArgAlaLys 260
 Db 781 AGCCGGACCCGCGACCCCTGTGACGAGGTGAAGAGCAGGCGGGAGGTGCGCGCAAG 840
 Oy 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleArgAlaLys 280
 Db 841 CTGAGAGGAGGAGCGCCAGCAGATACGCTCGAGCGCAGAGCCTTCCAGCGCGCTCAAG 900
 Oy 281 SerTrpPheGlnProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
 Db 901 AGCTGTTTCGAGCCCTGTGTGGAAGACATGACGCGCAGGCGGCGGCTGTGTGAGAAG 960
 Oy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 961 GTGCAGGCTGCGGTGGGACACGCGCGCCCTGTGTGCCAGCAGCAATTCAC 1011
 RESULT 9
 AX333278 1157 bp DNA linear PAT 09-JAN-2002
 LOCUS Sequence 3787 from Patent WO0194629.
 DEFINITION AX333278
 ACCESSION AX333278.1 GI:18123912
 VERSION
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1
 AUTHORS Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,
 Horrigan, S., Soppet, D. R. and Weaver, Z.
 TITLE Cancer gene determination and therapeutic screening using signature
 gene sets
 JOURNAL Patent: WO 0194629-A 3787 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 FEATURES
 source 1..1157
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 212 a 370 c 426 g 149 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,87e-79
 Score: 1572.00
 Percent Similarity: 99.37%
 Best Local Similarity: 99.05%
 Query Match: 99.18%
 Length: 1157
 Matches: 314
 Conservative: 1
 Mismatches: 2
 Indels: 0

DB: 6 Gaps: 0
 US-09-827-854-19 (1-317) x AX333278 (1-1157)
 Oy 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 62 ATGAAGTTCTGTGGGCTGCGGCTTGTGCTGTCATCTCTGCGAGATGCCAGGCCAAGGTG 121
 Oy 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnLys 40
 Db 122 GAGCAAGCGGTGAGAGACAGAGCCGAGCCGAGCTGGCCAGCAGAGCGAGTGGCAGAGC 181
 Oy 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTyrPvalGlnThr 60
 Db 182 GGCAGCGCTGGAGACCTGGCACTGGGTGCTTTTGGATTAACCTGGCTGGGTGCAGACA 241
 Oy 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 Db 242 CTGCTGACGAGGCGAGAGAGAGCTCTCAGCTCCCAAGTCAACCCAAAGACTGAGGGCG 301
 Oy 81 LeuMetAspGlnTrpMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
 Db 302 CTGATGGACGAGACCATGAAGAGTTGAAGGCTTCAAAATCGGAACCTGGAGAACTAG 361
 Oy 101 ThrProValAlaGlnGlnGlnArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
 Db 362 ACCCGGTAGCGAGAGAGAGCGCGGCGCAGGCTGTCCAAAGAGCTGCAGAGCGCCAGGCC 421
 Oy 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 Db 422 CGCGTGGCGCGGACATGAGAGAGAGCTGTGCGCGCGCTGTGTGAGTACCGCGGAGAGTG 481
 Oy 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnGlnAlaGlnArgLeuAlaSerHisLeuArg 160
 Db 482 CAGGCCATGCTCGGCGCAGACACCGAGAGCTGGGGTGGCGCTCGCTCCACCTCGCGC 541
 Oy 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 Db 542 AAGCTGGTAAAGCGGCTCTCTCGGATCCCGATGACCTGCAGAAAGCCCTGGCAGTAC 601
 Oy 181 GlnAlaGlyAlaArgGlnGlnGlnAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
 Db 602 CAGGCGGGGCGCGCGAGGCGCGCAGAGCGGCGCTCAGCGCCATCCCGAGCGCGCTGGG 661
 Oy 201 ProLeuValGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
 Db 662 CCCCTGGTGAAAGAGGCGCGGCTGGGGCGCCACATGTGGGCTCCCTGGCGCGCAGCGG 721
 Oy 221 LeuGlnGlnArgAlaGlnAlaATTPGlyGluArgLeuArgAlaArgMetGlnGlnMetGly 240
 Db 722 CTACAGAGAGGCGCGCCAGGCGCTGGGGGAGAGCGGCTGGCGCGGAGTGAAGATGGGC 781
 Oy 241 SerArgThrArgAspArgLeuAspGluValGlyGlnGlnAlaGlnValArgAlaLys 260
 Db 782 AGTCGAGCCCGCAGCAGCCCTGTGAGAGAGTGAAGAGAGAGTGGCGGAGTGGCGCCAG 841
 Oy 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleArgGlnArgLeuLys 280
 Db 842 CTGAGAGAGAGCGCCAGCAGATACGCTCAGAGCGAGGCTTCCAGGCGCGCTCAAG 901
 Oy 281 SerTrpPheGlnProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
 Db 902 AGCTGTTTCGAGCCCTGTGTGGAAGACATGACAGGCGCAGTGGCGGCTGTGTGAGAAG 961
 Oy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 962 GTGCAGGCTGCGGTGGGACACGCGCGCCCTGTGTGCCAGCAGCAATTCAC 1012
 RESULT 10
 AX409597 1157 bp DNA linear PAT 14-JUN-2002
 LOCUS Sequence 2244 from Patent WO0229103.
 DEFINITION AX409597
 ACCESSION

VERSION	AX409597.1	GI:21442302
KEYWORDS		
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	1 Alvaras,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G. Gene expression profiles in liver cancer Patent: WO 0229103-A 2244 11-Apr-2002;	
AUTHORS	Alvaras,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.	
TITLE	Gene expression profiles in liver cancer	
JOURNAL	Patent: WO 0229103-A 2244 11-Apr-2002;	
FEATURES	GENE LOGIC INC (US)	
SOURCE	Location/Qualifiers	
	1..1157	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/note="EMBL/Genbank Accession No. M12529"	
BASE COUNT	212 a	370 c 426 g 149 t
ORIGIN		
Alignment Scores:		
Pred. NO.:	1.87e-79	Length: 1157
Score:	1572.00	Matches: 314
Percent Similarity:	99.37%	Conservative: 1
Best Local Similarity:	99.05%	Mismatches: 2
Query Match:	99.18%	Indels: 0
DB:	6	Gaps: 0
US-09-827-854-19 (1-317) x AX409597 (1-1157)		
QY	1	MethylValLeuTTPAlaAlaLeuIleuValThrpheLeuAlaGlyCysGlnAlaIysVal 20
Db	62	ATGAAGGTTCTGTGGCGCTTCTGTCATCTCTGCGAGATGCCAGGCCAAGGTG 121
QY	21	GIUGlnAlaValAGIuThrGluProGluProGluIleuArgGlnGlnThrGluTPGInSer 40
Db	122	GAGCAAGCGGTGGAGACAGACC CGGACCCGACGCTGGCCAGCAGACCGAGTGGCAGACC 181
QY	41	GLYGlnArgTPGILuLeuAlaLeuGlyArPheThrpAspIryLeuArgTPValGlnThr 60
Db	182	GGCCAGCGCTGGAGACTGGACTGGGTGGCTTTGGGATTAACCTGGCGCTGGGTGCAGACA 241
QY	61	LeuSerGlnGlnValAGlnGlnGlnIleuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db	242	CTGTCTGAGCAGGTGGAGAGAGACGTCGTGACGTCCCAAGTACACCCAAAGATGAGGGCG 301
QY	81	LeuMetAspGluThrMetIysGlnLeuLeuAlaIryIysSerGlnLeuGlnGlnLeu 100
Db	302	CTGATGGACGAGACCATGAAAGATTGAAGGCCCTCAAAATCGGAACGGAGGAACAACCTG 361
QY	101	ThrpValAlaGlnGlnGluThrArgAlaArgLeuSerIysGlnLeuGlnAlaAlaGlnAla 120
Db	362	ACCCCGGTAGCGGAGAGACCGCGGGCACGGCTGTCCAAAGAGCTGCCAGACGGCGCAGGCC 421
QY	121	ArgLeuGlnAlaAspMetGluAspValCysGlyArGLeuValGlnIryArgGlyGlnVal 140
Db	422	CGGCTGGGCGCGGACATGAGAGACGTGTGGCGCGCCCTGTGGCAGTACCGCGGAGGTG 481
QY	141	GlnAlaMetLeuGlnSerThrGlnGlnIleuArgValArgLeuAlaSerHisIleuArg 160
Db	482	CAGGCACATGCTCGGCCAGAGCACCGAGACACTGGGGTGGCGCTCGCTCCACCTCGCC 541
QY	161	LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnIysArgLeuAlaValIryr 180
Db	542	AAGTGGCTAAGCGGCTCTCCGGATCCCGATGACCTGCAGAAAGCGCTGGCAGGTGTAC 601
QY	181	GlnAlaGlyAlaArgGlnGlyAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
Db	602	CAGCGCCGGGGCCCGCGAGGGCGCGGACCGGCTTCAGCCCAATCCGCGAGCCCTGGGG 661
QY	201	ProLeuValGlnGlnGlnArgValArgAlaAlaIhrValIeIysSerLeuAlaGlnIlePro 220
Db	662	CCCTGTGGTGAACAGGCGCGGTGGGGCGGCACCTGTGGGCTCCCTGGCGGGCCAGCCG 721

OY	221	LeungIngluarAlaGlnAlaTrrptLyGluaTrgleuAraArgmeSgIuImetGly	240
Db	722	CACAGAGACGGGGCCAGGCCCTGTGGGGCGAGCGCGCGCGCATGGATGAGAGATTGGGC	781
OY	241	SerArgThrArqAspArgleuAspGluValLysGluGlnValAlaGluValAraArgAlaLys	260
Db	782	AGTCGGACCCTGGCAGCCGCTGGACGAGGTGAAGGAGACAGTGCCGAGAGTGGCCCCAAG	841
OY	261	LeuglInglInlaGlnInlaGlnInleaArgleuGlnAlaGluAlaPheGlnAlaArgleuLys	280
Db	842	CTGGAGAGACAGACCCACAGAGATACGGCTGCAGAGCGCGAGGCCCTTCACAGCCCGCTCAAG	901
OY	281	SerTrpPheGluProLeuValGluLysMetGlnAraArgIntPrAlaGlyLeuValGluLys	300
Db	902	ACCTGGTTTCAGACCCCTGTGTGGAAACAATGACAGCCCGCAGTAGTGGCGCGCTGTGGAGAAG	961
OY	301	ValGlnAlaAlaValGlyThrSerIlaAlaProValProserAspAsnHis	317
Db	962	GTCAGAGCTCCGTCGGGACACAGCCCGCCCTGTGCCACGACACAATCAC	1012
RESULT	11		
LOCUS	115975	1157 bp	DNA Linear PAT 03-APR-1996
DEFINITION	Sequence 1 from patent US 5472858.		
ACCESSION	U15975		
VERSION	U15975.1 GI:1250883		
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1157)		
AUTHORS	Attie,A.D., Gretch,D.G., Sturley,S.L. and Beckage,N.E.		
TITLE	Production of recombinant proteins in insect larvae		
JOURNAL	Patent: US 5472858-A 1 05-Dec-1995;		
FEATURES	Location/Qualifiers		
source	1..1157 /organism="unknown"		
BASE COUNT	212 a 370 c 426 g 149 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.87e-79	Length:	1157
Score:	1572.00	Matches:	314
Percent Similarity:	99.37%	Conservative:	1
Best Local Similarity:	99.05%	Mismatches:	2
Query Match:	99.18%	Indels:	0
DB:	6 Gaps:	0	
US-09-827-854-19 (1-317) x U15975 (1-1157)			
OY	1	MettysValIleuTrpAlaAlaIleuLeuValThrpHeuAlaGlyCyGlnAlaLysVal	20
Db	62	ATGAAGGTTCTGTGGGCTGCGCTGTGTGTACATTCTTCGACAGATGCCAGGCCAAGTG	121
OY	21	GluGlnAlaValGluThrGluProGluProGluLeuAraGlnGlnThrGluTrpGlnSer	40
Db	122	GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTGGCGCCAGACAGACCGAGTGGAGAGC	181
OY	41	GlyGlnAraGTPGluLeuAlaLeuGlyVrPheTrpAspTrpLeuAraTrpValGlnThr	60
Db	182	GGCCAGCCCTGGGAACCTGGCACTGGGTGCTTTGGGATTTACTGGCGTGGGTGCAGCA	241
OY	61	LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuAraGla	80
Db	242	CTGTCTGACAGGTGCAGAGAGAGCTGCTCAAGTCCCAAGTACCCACAGAACCTGAGGCG	301
OY	81	LeuMeLaspGluThrMetLysGluLeuLysAlaTryLysSerGluLeuGluGluGlnLeu	100
Db	302	CTGAAGCAGACGACCATTAAGAGAGTTGAAGGCCCTACAAATGCGAACTGGAGAACACTG	361
OY	101	ThrProValAlaGluGluThrAraGlaAraGluSerLysGluLeuGlnAlaAlaGlnAla	120
Db	362	ACCCGCTAAGCGAGAGAACCGGGCAGCGCTGTCTCAAGAGAGCTGCAGACGGCCAGGCC	421

QY	121	ArgLeuGluValAlaAspMetGluAspValCysGlyArgLeuValGlnIrrArgGlyGluVal	140	/gene="APOE"	/note="APOE mRNA; G00-119-691"
Db	422	CGGCTGGGCGCGGACATGAGAGACGTGTGGGGCGCTGTGCATACCGCGGCGAGGTG	481	62..1015	/note="APOE"
QY	141	GlnAlaMetLeuGlyGlnSerThrGluGlnLeuArgValArgLeuAlaSerHisLeuArg	160	/gene="APOE"	/note="precursor"
Db	482	CAGCGCATGTCTGGCCAGACACCGAGAGACTCGGGTGGCGCTCCGCTCCACCTGGCCG	541	/codon_start=1	/product="apolipoprotein E"
QY	161	LysLeuArgGlnArgLeuLeuArgAspAlaAspSerLeuGlnIrrArgLeuAlaValTyr	180	/protein_id="AAB59518.1"	/db_xref="GI:178849"
Db	542	AAGCTGTGTAAGCGGCTCCCGCGATCCCGATGACCTGCGAAGCGCCCTGGCAGGTAC	601	/db_xref="GDB:G00-119-691"	/translation="MKLVNALLVTFPLACGQAKKEQAVETEPPELRQOTENQSGRW
QY	181	GlnAlaGlyAlaArgGluGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200	/db_xref="GDB:G00-119-691"	ELALGRFMDYLKWVOTLSQVOEELSSQVTELRALMDETKELKAYSELEDLTP
Db	602	CAGGCGGGGCGCGGAGGGCGCGGAGCGGGCTCAGGCCCATCCCGAGCGCCCTGGGG	661	KLKRLRPDDLOKRLAVYQAGARGSLVQYRGVEYQAMAGOSTEELRYSLASLR	VAETFRARLSKEIQTOARIGADMEVCGRLVQYRGVEYQAMAGOSTEELRYSLASLR
QY	201	ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro	220	OPLOERAOAWGSRILARMEGSSRRPDRIDEVKEQVAYRKLREDAQOIRIQAEAFQ	ARKLSFEPFLVEDMQRQWAGLVKQVQAAVGTSAAPVPSDNH"
Db	662	CCCTGTGTGAAACAGGGCCGCTGGCGCGCCACTGTGGGTCCCTGGCGGCGCAGCCG	721	62..115	/gene="APOE"
QY	221	LeuGlnGluArgAlaGlnAlaThrGlyGluArgLeuArgAlaArgMetGluGluMetGly	240	/note="APOE"	/note="G00-119-691"
Db	722	CTACAGAGCGGGCCAGGCTGGGGGAGCGGCTGGCGCGCGGATGAGAGATGGGC	781	116..1012	/gene="APOE"
QY	241	SerArgThrArgAspArgLeuAspGluValIrrGluGlnValAlaGluValArgAlaLys	260	/product="apolipoprotein E"	/note="G00-119-691"
Db	782	AGTCGGAGCCCGGACCCCTGGAGACGAGCTGAAGAGACAGTGGCGGAGGTGCCGCCAAG	841	370	/gene="APOE"
QY	261	LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys	280	/note="a in PHAE[112,178]; g in PHAE813; G00-119-691"	/replace="g"
Db	842	CTGGAGAGACGAGCCAGCAGATACGCTGCAGAGCGGAGGCTTCCAGAGGCCCTCAAG	901	569	/gene="APOE"
QY	281	SerTyrPheGluProLeuValGluAspMetGluArgGlnTyrPalaGlyLeuValGluLys	300	/note="c in PHAE[112,178]; g in PHAE813; G00-119-691"	/replace="g"
Db	902	AGCTGTGTGAGCCCTGGGTGGAGACATGACAGCGCAGTGGGCGGTGTGGAGAAAG	961	784	/gene="APOE"
QY	301	ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis	317	/note="t in PHAE[112,178]; c in PHAE813; G00-119-691"	/replace="c"
Db	962	GTCAGAGCTCCGCTGGGACACACGCGCCCTGTGCTCCAGCACATATCAC	1012	859	/gene="APOE"
RESULT 12				/note="g in PHAE[112,178]; a in PHAE813; G00-119-691"	/replace="a"
HUMAPOE				variation	
LOCUS				variation	
DEFINITION				variation	
ACCESSION				variation	
VERSION				variation	
KEYWORDS				variation	
SOURCE				variation	
ORGANISM				variation	
REFERENCE				variation	
AUTHORS				variation	
TITLE				variation	
JOURNAL				variation	
MEDLINE				variation	
PUBMED				variation	
FEATURES				variation	
source				variation	
1..1157				variation	
/organism="Homo sapiens"				variation	
/db_xref="taxon:9606"				variation	
/map="19q13.2"				variation	
/clone="PHAE[112,178,813]."				variation	
/sex="male"				variation	
/tissue_type="liver"				variation	
/dev_stage="57-year old"				variation	
1..1157				variation	
/gene="APOE"				variation	
<1..1157				variation	
gene				variation	
mrna				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna				variation	
linear				variation	
PRI 08-AUG-1995				variation	
complete cds.				variation	
1157 bp				variation	
mrna					

D	Bd		302	CtGATGGACGAGACCCTGAAGAGTTGAAGGCTTACAATTCCGAAGCTGGAGAACACTG	361
Oy			101	TThrProValAlaGluGluThrArgAlaArgLeuSerLysIleuGlnAlaAlaGlnAla	120
D	Bd		362	ACCCCGGTAGCGGAGAGACGCCGGCACGGCTGTCCAAAGACTGCAGACGCCGACGCC	421
Oy			121	ArgLeuGlyAlaAspMetCyluAspValLysGlyArgLeuValGlnThrArgGlyVal	140
D	Bd		422	CGGCTGGGGCGGACATGGAGAGGTGTGGGGCGGCTGTGTCAATGACGGCGGAGGGTG	481
Oy			141	GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg	160
D	Bd		482	CAGGCATCTCTGGCCAGACACCGAGAGCTGGGGGTGGCTGGCTCCACCTGGCC	541
Oy			161	LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
D	Bd		542	AAGCTCGTAACCGGCTCTCCGCGATCCCGATGACCTGCACAACCGCTGCAGAGTAC	601
Oy			181	GlnAlaGlyAlaArgGluGluAlaGluAlaArgGlyLeuSerAlaIleArgLysArgLeuGly	200
D	Bd		602	CAGCGCGGGCCCCGCGAGGGCCCCGACGCCGCGCTCAGGCCCATCCGACGCGCTGGGG	661
Oy			201	ProLeuValGluGluGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlyLeuPro	220
D	Bd		662	CCCCTGGTGAACAGGGCCGCTGGGGCGCCGACAGTGGCTCCCTGGCCGGCACGGC	721
Oy			221	LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly	240
D	Bd		722	CTACAGAGACGGGGCCCCAGGCTTGGGGCGACGGCGCTGGCGGATGGGC	781
Oy			241	SerArgTrpArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys	260
D	Bd		782	ACTCGGACCCCGGACCGCCTGGACAGAGGAGAAGACAGTAGTGGCGAGTGGCGCCAG	841
Oy			261	LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys	280
D	Bd		842	CTGGAGAGACAGCCAGCCAGATAGCCCTGCAGGCCGAGCCCTTCAGGCCGCTCAAG	901
Oy			281	SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys	300
D	Bd		902	ACTGGTTCGAGCCCTGTGTGGAGACATGACAGCGCAGTGGCGGGCTGTGTGGAGAA	961
Oy			301	ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis	317
D	Bd		962	GTGACGGCTGCGGTGGGCACACGCCCGCCCTGTGCTCCAGCAGCAATCATC	1012
RESULT 13					
BD004279					
LOCUS BD004279 1156 bp DNA linear PAT 31-JAN-2002					
DEFINITION Apo E humanized mammal.					
ACCESSION BD004279.1 GI:18632240					
VERSION BD004279.1					
KEYWORDS JP 2001017028-A/3.					
SOURCE Homo sapiens.					
ORGANISM Homo sapiens.					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.					
REFERENCE 1 (bases 1 to 1156)					
AUTHORS Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.					
TITLE Apo E humanized mammal					
JOURNAL Patent: JP 2001017028-A 3 23-JAN-2001;					
COMMENT MITSUBISHI CHEMICAL CORP					
OS Homo sapiens (human)					
PN JP 2001017028-A/3					
PD 23-JAN-2001					
PF 28-APR-2000 JP 2000128919					
PR					
PI SHINOBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA PC					
AOIK67/027,A6IK45/00,A6IP25/28,A6IP43/00,C12N5/10,PC					
C12N1/09,/CO7K14/775,					
PC (C12N5/10,C12R1:91),C12N1/00,C12N15/00,(C12N5/00,C12R1:91) CC					

	FEATURES	source	FH FT CDS	Key Location/Qualifiers Location/Qualifiers (61)..(1011).
				1..1156 /organism="Homo sapiens" /db_xref="taxon:9606"
	BASE COUNT	208 a 369 c 432 g 147 t		
	ORIGIN			
	Alignment Scores:			
	Pred. No. :	2.74e-79	Length:	1156
	Score:	1569.00	Matches:	315
	Percent Similarity:	99.68%	Conservative:	1
	Best Local Similarity:	99.37%	Mismatches:	1
	Query Match:	98.99%	Indels:	0
	DB:	6	Gaps:	0
	US-09-827-854-19 (1-317) x BD004279 (1-1156)			
Oy	1 MetLysValLeuTrpAlaIalaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal			20
Db	61 ATGAAGATTCTGTGGGCTCCGTTCCTGGTCAATTCCTCGCAGACATGCCAAGAAGTG			122
Oy	21 GlnGlnAlaValaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer			40
Db	121 GAGCAAGCCGTGGAAACAAGAGCCGGAAGCCGAGCTGCACCAGACAGCCACTGGCAGAGC			180
Oy	41 GlyValArgTrpGluLeuAlaLeuGlyLyrPheTrpAspPyrLeuArgTrpValGlnThr			60
Db	181 GGCCAGCGCTGGAACTGGCACTGGTGGCTTTGGATTACCTGCGCTGGCTGACACA			240
Oy	61 LeuSerGluGlnValaGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla			80
Db	241 CTGTCTGACACAGTCAGAGAGAGCTGCTCACGCTCCAGCTACACCAGAACCTAGAGCGC			300
Oy	81 LeuMetAspGluThrMetLysGluLeuLysAlaTryLysSerGluLeuGlnGlnLeu			100
Db	301 CTGATGAGACGAACCATGAAGAGTTGAAGGCTTACCAATGCGAAGCTGAGAACCACTG			360
Oy	101 ThrProValaLaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla			120
Db	361 ACCCGGTCGGCGAAGAGACGGCGGCACGGCTGTCCAAGAGAGCTGACGGCGGCAGGCC			420
Oy	121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTryArgGlyGluVal			140
Db	421 CGGCTGGGCGCAGATGAGAGACTGGCGCGCCCTGGTGTCAAGTACCGCGGAGAGTG			480
Oy	141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValaArgLeuAlaSerHisLeuArg			160
Db	481 CAGGCCATGCTGTGGCGACAGACACCGAAGAGACTGGCGGCTGCGCTCCCACTGGCC			540
Oy	161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr			180
Db	541 AAGCTGCGTAAGCGGCTCTCCGCGATCCCGATGACCTGCAGAAAGCGCTGGCAGTGNAC			600
Oy	181 GlnAlaGlyAlaArgGluGlyAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly			200
Db	601 CAGGCGGGGCGCGGAGGCGCGGAGGCGGCGCTTCACAGCCATCCGAGACGCTGGGG			660
Oy	201 ProLeuValaGluGlnGlyArgValaArgAlaIleThrValaGlySerLeuAlaGlyGlnPro			220
Db	661 CCCCTGGTGGAAACAGAGCGCGCTGGGGCCCCCACTGGGCTCCCTGCGCGGCAAGCG			720
Oy	221 LeuGlnGluArgAlaGlnAlaIleTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly			240
Db	721 CTCACGAGGAGCGGCGCCACAGGCTGGGGCGAAGCGCTGCGCCGCGGATGAGAGATGGGC			780
Oy	241 SerArgTrpArgAspArgLeuAspGluValaLysGluGlnValaAlaGluValaArgAlaLys			260
Db	781 ACCCGGACCGGACCGGCTGACAGAGGTGAAGAGCGAGGTGGGAGAGTCCGCGCCAAAG			840
Oy	261 LeuGlnGluGlnAlaGlnGlnIleArgArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys			280

Db 841 CTGAGAGAGAGAGAGAGAGATACCGCTGACAGCGCCCTTCCAGCGCCCTTCAG 900
QY 281 SerTrrPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 901 AGCTGGTTCAGAGCCCTGGTGGAGAACATGACAGCGCCAGTGGGGCGGCTGGTGGAGAG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTCCAGAGCTCCGTGGGACACAGCGCCGCTGTGGCCAGCAGCAATATC 1011
RESULT 14
MFAPOE 1178 bp mRNA linear PRI 31-MAR-1995
LOCUS MFAPOE
DEFINITION Monkey mRNA for apolipoprotein E.
ACCESSION X13887
VERSION X13887.1 GI:38054
KEYWORDS apolipoprotein; apolipoprotein E.
SOURCE Macaca fascicularis.
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecoidea; Macaca.
REFERENCE 1 (bases 1 to 1178)
AUTHORS Marotti K.R.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1988) Marotti K.R., The Upjohn Co, 7242 209 713,
301 Henrietta Street, Kalamazoo, MI 49008
2 (bases 1 to 1178)
AUTHORS Marotti K.R., Whitted, B.E., Castle, C.K., Polltes, H.G. and
Melchior, G.W.
TITLE Nucleotide sequence of the cynomolgus monkey apolipoprotein E cDNA
JOURNAL Nucleic Acids Res. 17 (4), 1778 (1989)
MEDLINE 89160349
PUBMED 2922300
COMMENT Data kindly reviewed (20-Mar-1989) by Marotti K.R.
FEATURES
source
1..1178
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="1"
/tissue_type="liver"
/clone_lib="lambda gt11"
83..1036
/codon_start=1
/product="proapolipoprotein E (AA 1-18 to 299)"
/protein_id="CAA32092.1"
/db_xref="GI:38054"
/db_xref="SWISS-PROT:P10517"
/translation="MKVLAALIVTFVLACGQAKVQEPVETPEPELRQAGGSGQPM
EVALGRFMDYLRVQTLSEVOEELSPQVTELTIMDETAKELKAYSELEQISP
VAETRRARLSKEVLOAARLGADEMDVRSRLVOYSEVOAMLGSGSEELRRLASHLR
KLRKRLRDADLQRLRIAYOAGARGARSGVSLRERGPVDEGGRVRAATVGSLSAS
QPLQERAAQALGRLRLARMEKMSRTDRIDVEKVAEYRAKLEQADQISLQAEKFO
ARKLSMEPELVDMQDMQWAGLVEKVAQAASTAPVPIDNH"
sig_peptide
mat_peptide
misc_feature
BASE COUNT 217 a 373 c 426 g 162 t
ORIGIN
Alignment Scores:
Pred. No.: 1.7e-72 Length: 1178
Score: 1447.00 Matches: 295
Percent Similarity: 94.018 Conservative: 3
Best Local Similarity: 93.068 Mismatches: 19
Query Match: 91.298 Indels: 0
DB: 9 Gaps: 0
US-09-827-854-19 (1-317) x MFAPOE (1-1178)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20

Db 83 ATGACAGCTTCTGGGCTGCTGGCTGCTGCATCTCTGACAGATCCAGGCCAAGGTG 142
QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnTrpGlnSer 40
Db 143 GAGCAACCGGTGGAGCCAGACGAGACCCGACCTTCGACAGAGCTGAGCGCCAGAGC 202
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspPylrLeuArgTrpValGlnThr 60
Db 203 GGGCAGCCCTGGGAGCTGGGACCTGGGCTGGTCTTTGGATTACTGGCTGGTGGACACA 262
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db 263 CTGTCTGACAGAGGTGACAGAGAGAGAGCTGCTGACGCCAGAGCTACCCAGAACTGACAGC 322
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaThrLysSerGluLeuGlnGlnLeu 100
Db 323 CTGATGACGAGACCATGAGAGAGTGAAGCCCTTACAAATCGAAGCTGGAGAAACAGCTG 382
QY 101 ThrProValAlaGlnGlnGlnThrArgAlaArgLeuSerLysGlnGlnAlaAlaGlnAla 120
Db 383 AGCCCGGTGGCGAGAGAGACCGGGCAGGCTGTCCAGAGACTGCGAGCGCGCCAGGCC 442
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyVal 140
Db 443 CGCTGGGTGCTCCGACATGAGAGAGAGAGCTGCGACCGGCTGTGTCAGTCCGACAGAGGTG 502
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
Db 503 CAGGCGATGCTGGGGCCAGACTACGAGAGAGCTGCGGGCGGCTGCTCCACCTGCGC 562
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 563 AACCTGCGCAGAGCGGCTCTCCGCGATGCTGATGACTGTCAGAAAGCCCTGCGCAGTGTAT 622
QY 181 GlnAlaGlyAlaArgGlnGlyAlaGlyArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 623 CAGCGCGGGCGCGCGAGGGCGCGAGAGCGGGGTGACGCCCATCCGCGAGCGCGCTGGGA 682
QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 683 CCCCTGGTGGAGAGGGCGCGCTGCGGGCGCGCCACTGTGGGCTCCCTGGCGCACCGCGG 742
QY 221 LeuGlnGlnArgAlaGlnAlaThrGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
Db 743 CTTCAGAGAGGGCGCCAGAGCTTGGGTGAGCGGCTTCGCGCAGGATGAGAGATGGGC 802
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
Db 803 AGCCGAGACCGCGACCGCTGCGAGAGAGTGAAGAGAGAGAGTGGCGGAGTGGCGCCAG 862
QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgGlnGlnAlaGlnAlaIleGlnAlaArgLeuLys 280
Db 863 CTGAGAGAACAGGCCAGAGATMACCTGACAGCGGAGGCTTCCAGAGCGCCCTTCAG 922
QY 281 SerTrrPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 923 AGCTGGTTCAGAGCCCTGGTGGAGAACATGACAGCGCCAGTGGGCTGGTGGAGAGAG 962
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 983 GTCCAGAGCTCCGTGGGCGCCAGCAGCGCCGCTGTGGCCATGCAATATC 1033
RESULT 15
AF261279 5491 bp DNA linear PRI 27-OCT-2000
LOCUS AF261279
DEFINITION Homo sapiens apolipoprotein-E gene, complete cds.
ACCESSION AF261279
VERSION AF261279.1 GI:11034800
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 5491)
AUTHORS Nickerson,D.A., Taylor,S.L., Fullerton,S.M., Weiss,K.M.,
Clark,A.G., Stengard,J.H., Salomaa,V., Boerwinkle,E. and Sing,C.F.
TITLE Sequence diversity and large-scale typing of SNPs in the human
apolipoprotein E gene
JOURNAL Genome Res. 10 (10), 1532-1545 (2000)
MEDLINE 20499366
PUBMED 11042151
REFERENCE 2 (bases 1 to 5491)
AUTHORS Nickerson,D.A.
TITLE Direct Submission
JOURNAL Submitted (27-Apr-2000) Department of Molecular Biotechnology,
University of Washington, Box 357730, Seattle, WA 98195, USA
FEATURES
source 1. 5491
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.2"
/contig="108"
repeat_region /note="putative"
/rpt_family="MIR"
/rpt_type="dispersed"
73
variation /frequency="0.01"
/replace="t"
/note="putative"
/rpt_family="MIR"
/rpt_type="dispersed"
308
variation /frequency="0.01"
/replace="t"
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
471
variation /frequency="0.01"
/replace="g"
545
variation /frequency="0.01"
/replace="t"
560
variation /frequency="0.22"
/replace="t"
624
variation /frequency="0.07"
/replace="c"
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
832
variation /frequency="0.45"
/replace="t"
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
1163
variation /frequency="0.35"
/replace="c"
1522
variation /frequency="0.01"
/replace="a"
1575
variation /frequency="0.01"
/replace="t"
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
1878
variation /frequency="0.06"
/replace="c"
/note="APOE"
/codon_start=1
/product="apolipoprotein-E"
/protein_id="AAG27089.1"
/db_xref="GI:11034801"
/translation="MKVLMALLVTFLAGCOAKVEQAVETPEPELRROTEMOGQRM
ELALGPRWDYLRVQYRLSROYQFELISSVYTOPLRALMDETMKELAKYSELEQLTP
VABETRLAKSLKELQAOARLGADMEDYCGRLVYRGQVOMLQOSTEELRYLASHLR
KLKRLRLDQADDLQKRLAVYQAGARBEAGELSAIRRLGLPVEQGVRAVAGSLAG
QPLQERQADGGERLRAKREMGSGRTDRDLDEKQVQAEVRAKLEQAOQIRLQAEAFQ
ARLSWEPEPLVEDMRQMAGLVEKQVAAVTSAAVPSPDNH"
1998
variation /frequency="0.10"
/replace="a"
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
2440
variation /frequency="0.21"
/replace="a"
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
2907
variation /frequency="0.01"
/replace="g"
3106
variation /frequency="0.01"
/replace="c"
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
3673
variation /frequency="0.01"
/replace="g"
3701
variation /frequency="0.01"
/replace="ct"
3937
variation /frequency="0.14"
/replace="c"
4036
variation /frequency="0.01"
/replace="t"
4075
variation /frequency="0.07"
/replace="t"
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
5065
variation /frequency="0.03"
/replace="gg"
5229
variation /frequency="0.07"
/replace="t"
5229
variation /frequency="0.40"
/replace="g"
5230
variation /frequency="0.13"
/replace="a"
5361
variation /frequency="0.06"
/replace="c"
/note="APOE"
/codon_start=1
BASE COUNT 1029 a 1663 c 1603 g 1196 t
ORIGIN

Alignment Scores:

Pred. No.:	8.97e-69	Length:	54
Score:	1392.80	Matches:	303
Percent Similarity:	60.80%	Conservative:	1
Best Local Similarity:	60.60%	Mismatches:	2
Query Match:	87.85%	Indels:	194
DB:	9	Gaps:	1

US-09-827-854-19 (1-317) x AF261279 (1-5491)

[illegible]

Db	3903	GCAGGCCCGGCTGGAGCCGGACATGGAGAGCACTGTGGCCGCCCTGGTGCAGTTACCCGGG	3962
OY	138	YGLVVALGlnAlaMetLeuGlnGlnSerThrGlnGlnLeuAraValAraGluAlaSerHis	158
Db	3963	CGAAGTGCAGGCCATGCTCTCGGCCAGAGACACCGAGGAGCTGCGGGTGCCTCGCTCCCA	4022
OY	158	SLenAArgLysLeuAraGlnGlnAraGluLeuAraGspAraAraPheLeuGlnLysAraGluAl	178
Db	4023	CTCTCGCAAGCTGGGTAAAGGCGGTCTCTCCGCGCATGCGCATGCAACCTGCAGAACGCTGGC	4082
OY	178	AAValTyrGlnAlaGlyAraArgGlnGlnGlyAlaGluAraGlyLeuSerAlaLeaArgLuar	198
Db	4083	AGTGTACCAAGCCCGGGGCCGCCAAGGGCCGCCAGCCGGGCTCAAGCCATCCGGAGCG	4142
OY	198	GLenGlyProLeuValGlnGlnGlnLysAraValAraGlnAlaThrValGlySerLeuAlaG	218
Db	4143	CCTGGGGGCCCTGGTGTGAACAGGGCCGCGTGGGGCCGCCCATGTGGGCTCCCTGGCCGG	4202
OY	218	YGlnProLeuGlnGlnLuarGlnAlaGlnAlaThrProGlyGlnAraGluAraGlnMetGlnG	238
Db	4203	CCACGCCCTCAAGAGAGCGGGGCCCGCAGGCTCTGGGGCGAGCCGCTGCAGCGCGGATGAAGA	4262
OY	238	UMetGlySerAraGhrThrArgPheLeuAraGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnG	258
Db	4263	GATGGGAGCGCCGACCCGCGACCGCTGTGGACGAGGTGAAGAGAGCGAGTGGCGAGCTCG	4322
OY	258	GAlaLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnG	278
Db	4323	CGCCAAAGCTGAGAGAGAGAGGCCCGACAGATACCCCTGCAGAGGCCAGAGCCTTCCAGGCCG	4382
OY	278	GLenLysSerTyrPheGlnProLeuValGlnAraPheMetGlnAraGlnTyrPalaGlyLeuA	298
Db	4383	CCCTCAAGAGCTGGTTTCAGAGCCCTCTGGTGTGAACATAGCAGGCCCATGTGGCCGGGCTGT	4442
OY	298	IGlnLysValGlnAlaLeuValGlyThrSerAlaAraProValProSerAraPheHis	317
Db	4443	GGAAAGAGTGTAGGCTCTCCGTGGGCACACAGCGCCGCCCTGTGGCCAGAGCAATCAC	4500

Search completed: March 14, 2003, 17:32:49
Job time : 2313.23 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:08:17; Search time 178.728 Seconds
(without alignments)
3994.237 Million cell updates/sec

Title: US-09-827-854-19

Perfect score: 1585

Sequence: 1 MKVLMAALLVFLAGQAKV.....VEKVOAAGVTSAAVPSPDNH 317

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=yld
-O=/cgr2.1/USFTO.spool/US0827854.r/unat.11032003_101609_27476/app.query.fasta_1.3576
-DB=N.Geneseq-101002 -QPM=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cd1
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US0827854.@CGN.1.1201.ernalat.11032003.101609.27476 -NCPU=6 -ICPU=3
-NO_XLPEXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEDOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database: N.Geneseq.101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001C.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1585	100.0	1156	24	AAD22052	Human apolipoprotein E
2	1581	99.7	954	24	AAD26035	Human apolipoprotein E
3	1581	99.7	1110	7	AAN60409	Human apolipoprotein E
4	1581	99.7	1147	22	ABA83113	Human apolipoprotein E
5	1581	99.7	1156	22	AAE84315	Human apolipoprotein E
6	1581	99.7	1156	24	AAD22048	Human apolipoprotein E
7	1573	99.2	1156	22	AAE84314	Human apolipoprotein E
8	1573	99.2	1156	24	AAD22049	Human apolipoprotein E
9	1573	99.2	1156	24	AAD22051	Human apolipoprotein E
10	1572	99.2	1157	17	AAE06957	Human apolipoprotein E
11	1572	99.2	1157	24	ABN95746	Gene #2244 used to
12	1572	99.2	1157	24	ABN64514	Human benign prot
13	1572	99.2	1157	24	ABN64510	Lung cancer relate
14	1571	99.1	1110	6	AAN50450	Sequence encoding
15	1569	99.0	1156	22	AAE84316	Human ApoE coding
16	1569	99.0	1156	24	AAD22047	Human apolipoprotein
17	1566	98.8	1156	24	AAD22050	Human apolipoprotein
18	1477.5	92.0	1279	22	AAE22437	Human apoE encodin
19	1458.5	92.0	1107	19	AAE75756	Human apolipoprotein
20	1392.5	87.9	9360	24	ABL31915	Human ApoE genomic
21	1392.5	87.9	10716	24	AAD26034	Human apolipoprotein
22	1387	87.5	3805	20	AAE09534	Human Apo E genom
23	1387	87.5	3805	20	AAE09526	Human Apo E genom
24	1365.5	86.2	10716	24	AAD26108	Human apolipoprotein
25	1159.5	73.2	965	24	AAD32081	Bovine ApoE gene.
26	1159.5	73.2	5617	24	AAD32077	Human alpha-1-anti
27	1159.5	73.2	6026	24	AAD32075	Human albumin prom
28	1134	71.5	1126	19	AAE29159	Nucleotide sequenc
29	981	61.9	936	15	AAE69101	ApoA4L2 protease
30	981	61.9	936	17	AAE18070	Human ApoA4L2 CN
31	980	61.8	660	18	AAE69792	Partial human apol
32	943	58.9	597	17	AAE18068	Human ApoA4 CNNA.
33	934	58.9	597	15	AAE69099	ApoA4L protease cd
34	880	55.5	1381	22	AAE22673	Human cDNA encodin
35	867.5	54.7	786	21	AAE18114	Lung cancer associ
36	674	42.5	600	20	AAE89595	EST clone Cpl47.
37	651	41.1	407	24	ABK34238	Human cDNA for nov
38	625	39.4	478	24	ABN95594	Gene #2492 used to
39	625	39.4	478	24	ABE62659	Colon adenocarcino
40	625	39.4	478	24	ABE67340	Thyroid cancer rel
41	594	37.5	459	22	AAE503049	Human diagnostic a
42	510	32.2	330	12	AAQ11980	Human apolipoprotein
43	504	31.8	405	21	AAE02139	Human secreted pro
44	485	30.6	345	22	AAE98479	Human EST-derived
45	415	26.2	260	21	AAE40342	Human apolipoprotein

ALIGNMENTS

RESULT 1
AAD22052
ID AAD22052 standard; DNA: 1156 BP.

AC AAD22052:
DT 12-FEB-2002 (first entry)

DE Human apolipoprotein E (apoE) allele, apoE2** DNA.

KW Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;

KX hypertriglyceridaemia; low density lipoprotein; LDL; ds.

OS Homo sapiens.

XX key Location/Qualifiers
FH 61..1014
FT CDS /tag= a

FT sig_peptide /product= "Human apoe allele, apoe2**"
 FT 61..114 /*tag= b
 FT mat_peptide 115..1011
 FT /*tag= c
 FT /product= "Mature human apoe allele, apoe2**"
 XX W0200177136-A1.
 XX 18-OCT-2001.
 XX
 XX 06-APR-2001; 2001WO-US11358.
 XX
 XX 06-APR-2000; 2000US-0544386.
 PR 04-OCT-2000; 2000US-0679088.
 PR 05-APR-2001; 2001US-0827854.
 XX
 PA (KOSP-) KOS PHARM INC.
 PA (UYBO-) UNIV BOSTON.
 XX
 PI Zannis VI, Kyreos KE;
 XX
 DR WPI: 2002-010885/01.
 DR P-PDB: AAE13298.
 XX
 PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
 PT cholesterol, delaying the onset of or treating atherosclerosis in
 PT mammal, without inducing hypertriglyceridemia -
 XX
 PS Claim 14; Page 83; 91pp; English.
 XX
 CC The present sequence is a human apolipoprotein E (apoE)
 CC allele, apoe2** DNA. The apoE lipoproteins are useful for
 CC lowering cholesterol, delaying the onset of atherosclerosis,
 CC treating or regressing atherosclerosis without inducing
 CC hypertriglyceridaemia, in a mammal lacking an endogenous,
 CC normally functioning apoE gene or low density lipoprotein (LDL)
 CC receptor or is at risk for developing atherosclerosis due to
 CC accumulation of lipoprotein remnants in the bloodstream or having
 CC a defect in remnant removal.
 CC
 XX
 SQ Sequence 1156 BP; 207 A; 369 C; 432 G; 148 T; 0 other;

 Alignment Scores:
 Pred. No.: 4.32e-111 Length: 1156
 Score: 1585.00 Matches: 317
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

 US-09-827-854-19 (1-317) x AAD22052 (1-1156)

 QY 1 MetlyValleuTrrPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 61 ATGTAAGGTTCTGTGGGCTCTCTCTGTCATTCCTGGAGATGCCACCAAGGTG 120
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnThrGluTrpGlnSer 40
 Db 121 GAGCAAGCGGTGAGACAGACAGCCGAGCCGAGCTGGCCAGCAGACCGAATGGCAGAGC 180
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 Db 181 GGCAAGCGGTGGAATGCGACTGCTTGGATTACCTGCGCTGGTGCAGACA 240
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 Db 241 CTGCTTAGCAGGTGACGAGAGACCTGCTCACTCCAGAGTCAACCAAGTGAAGGCG 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 Db 301 CTGATGACGACGACCATGAAAGAGTTGAAGGCTTACAAATCGAATGAGAGCAACTG 360

QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
 Db 361 ACCCGGTGGCGGAGAGACCGGGGCAAGGCTGTCCAAAGAGTGCAGAGCGCCAGAGCC 420
 QY 121 ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGluVal 140
 Db 421 CGGCTGGGCGGACATGAGAGACGTGGCGGCGCTGGTGCATGACCGGGGAGGTG 480
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 481 CAGGCCATGCTGGGCCAGAGACAGCAGAGAGCTGGGGTGGCGCTCCGCTCCACTGCGC 540
 QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 Db 541 AAGCTGGCTCAGCGGCTCTCCGCAATGCCATGACCTCGAAGAGCCCTGGCAGTGTAC 600
 QY 181 GlnAlaGlyAlaArgGluGluValArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db 601 CAGGCGGGGCGCCGAGGGCGCGAGCGCGGCTTCAGCGCATCCGAGCGCGCTGGG 660
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 661 CCCCTGGTGGAAACAGGGCGCGGTGCGGGCGCCCACTGTGGGCTCCGTGGCGGCAAGCG 720
 QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
 Db 721 CTACAGAGAGCGGCGCCAGGCTGCGGCGAGCGGCTGCGCGCGGATGAGAGATGGCG 780
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 Db 781 AGCCGAGACCCGCGACCCCTGAGAGGTGAAGAGCAAGGAGGAGGAGGCGGCAAG 840
 QY 261 LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 Db 841 CTGAGAGGAGCGGCCAGCAGATACGCTGAGGCGCAGGGCTTCAGAGCGCCGCTCAG 900
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 Db 901 AGCTGTGTCGAGCCCTCGGTGGAGACATGACAGGCCAGTGGCGGCGGTGTGGAGAG 960
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 961 GTGCAGGCTGCGTGGACACAGCGCCGCCCTGTGCCAGCAATAC 1011

 RESULT 2
 AAD26035
 ID AAD26035 standard; cDNA; 954 BP.
 XX
 AC AAD26035;
 XX
 XX 26-MAR-2002 (first entry)
 DT
 XX
 DE Human apolipoprotein E (APOE) cDNA.
 XX
 KW Human: anti-lipemic; neuroprotective; noctropic; genetic variant; APOE;
 KW apolipoprotein E; haplotyping; familial dysbetalipoproteinemia; therapy;
 KW genotyping; type III hyperlipoproteinemia; Alzheimer's disease; SNP;
 KW atherosclerosis; single nucleotide polymorphism; chromosome 19q13.2; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..954
 FT /*tag= a
 FT /product= "Human APOE protein"
 FT replace (13, C)
 FT variation
 FT /*tag= a
 FT /standard_name= "Single nucleotide polymorphism (SNP)"
 FT replace (31, G)
 FT /*tag= b
 FT /strand.name= "Single nucleotide polymorphism (SNP)"
 FT replace (364, A)
 FT /*tag= c
 FT variation

PN AUB547513-A.
 XX 24-APR-1986.
 XX 17-SEP-1985; 85AU-0047513.
 XX 11-JUN-1985; 85JP-0126989.
 XX (MITU) MITSUBISHI CHEM IND KK.
 PA WPI: 1986-150217/24.
 DR P-PSDB; AAP60507.
 PT New DNA sequence coding for human apolipoprotein-E - and
 PS expression vectors and transformed cells contg. it.
 XX Disclosure; Fig 2; 45pp; English.
 CC The encoded protein is used to treat subjects who are deficient in
 CC apolipoprotein-E (or who produce abnormal forms of this molecule)
 CC and therefore are likely to suffer from hyperlipidemia, resulting in
 CC arteriosclerosis. It can also be used to raise antisera for
 CC detecting the protein deficiency or production of abnormal forms.
 XX Sequence 1110 BP; 198 A; 354 C; 415 G; 143 T; 0 other;

Alignment Scores:

Pred. No.: 8,28e-111 Length: 1110
 Score: 1581.00 Matches: 316
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.68% Mismatches: 0
 Query Match: 99.75% Indels: 0
 DB: Gaps: 0

US-09-827-854-19 (1-317) x AAN60409 (1-1110)

QY 1 MetysValIeuTrrpAlaIalaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 15 ATGAAAGGTCTGTGGTGGCTGTCTGTGTCATCTCTGGCAGAGATGCCAGGCCAAGGTG 74
 QY 21 GlnAlaIalaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 75 GAGCAAGCGGGAGACAGACGCCGAGCTGCGCCCGACGACCGAGTGGCAGAGC 134
 QY 41 GlyIlnaTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB 135 GGCACGGCTGGAACTGGACCTGGTGGATTACCTGGCGCTGGTGGTGCAGACA 194
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerGlnValThrGlnGluLeuArgAla 80
 DB 195 CTGTCTGAGCGAGTGCAGAGAGAGCTCTCAGCTCCAGAGTCCAGGAAGTGAAGGCG 254
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 DB 255 CTGATGACAGACCATGAAAGAGCTTAAAGCTTACAATCGGAGCTGGAGAACTG 314
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 315 ACCCGGTGGGGAGAGACCGGGGCAAGGCTGTCCAAGAGAGTGCAGAGGCGGAGGCC 374
 QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
 DB 375 CGGCTGGGGCCCGCAGAGGAGAGCTGTGGCGCTGTGTGCTAGTACCGCGGAGAGT 434
 QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 435 CAGGCCATGCTCGGCCAGAGACACCGAGAGCTGGGGTGGCTCCGCTCCACCTCGGC 494
 QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 DB 495 AAGCTGGTAAGCGGCTCTCCGCGATGCCGATACCTGCAGAAAGCGCTGGCAGTGTAC 554
 QY 181 GlnAlaGlyAlaArgGluGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200

DB 555 CAGCCCGGGGCCCGAGGGGCCGAGCGCCCTCAGCGCCATCCCGAGCGCTGGG 614
 QY 201 ProLeuValGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
 DB 615 CCCCTGGTGAACAGAGGGCGGCTGGGGCCGCACTGTGGCTCTCGTGGCGGACCG 674
 QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
 DB 675 CTACAGAGAGCGGGCCAGGCTGGGGCGAGCGCTGGCGCGCGGATGAGAGATGGGC 734
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
 DB 735 AGCCGGACCCCGCAGCCCTGGAGAGGTGAAGAGAGAGGGGGAGGTGGCGCCAAAG 794
 QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLys 280
 DB 795 CTGCAGACAGCAGGCCCGCAGCATACGCTCGAGGCCAGGCCCTTCCAGGCCGCTCAAG 854
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 DB 855 AGCTGGTTGAGCCCTGTGGTGAAGACATGACAGCGCCAGTGGCGGGCTGTGGAGAG 914
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 915 GTGCAGGCTGCCGTGGGACACAGCGCCGCTGTGCCACGACATAC 965
 RESULT 4
 ABA83113
 ID ABA83113 standard; DNA; 1147 BP.
 AC ABA83113;
 XX 08-FEB-2002 (first entry)
 DE Apolipoprotein E ovarian tumour marker gene, SEQ ID NO:63.
 XX Ovarian tumour marker gene; human; overexpression; upregulation;
 KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;
 KW identification; serous cystadenoma; borderline serous tumour;
 KW serous cystadenocarcinoma; mucinous cystadenocarcinoma;
 KW mucinous cystadenoma; borderline mucinous tumour; endometrioid carcinoma;
 KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;
 KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;
 KW immune response pathway; cell proliferation regulation; protein folding;
 KW membrane localised; secreted; therapeutic target; cytostatic;
 KW gene therapy; vaccine; ds.
 XX Homo sapiens.
 OS
 XX
 PN WO200175177-A2.
 PD 11-OCT-2001.
 PF 03-APR-2001; 2001WO-US10947.
 XX
 PR 03-APR-2000; 2000US-194336P.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PT Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;
 DR WPI: 2001-626450/72.
 DR P-PSDB; ABB50287.
 XX
 PT Detecting and identifying ovarian tumor, identifying increased risk for
 PT developing ovarian cancer, and determining effectiveness of ovarian
 PT cancer treatment, by measuring expression level of ovarian tumor marker
 PT gene -
 PS Claim 23; Page 105-106; 140pp; English.
 XX The invention relates to methods for diagnosing and prognosing ovarian

tumours in an individual via the detection and measurement of the expression of ovarian tumour marker genes (ABA83081-ABA83122, ABA83180, ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179, ABA83181 and ABA83183). The methods of the invention are useful for detecting an ovarian tumour in a patient, for identifying an individual at increased risk for developing ovarian cancer, in prognostic tests for assessing the relative severity of ovarian cancer, in tests for monitoring a patient in remission from ovarian cancer and in tests for monitoring disease status in a patient being treated for ovarian cancer. The methods can additionally be used to identify a particular tumour as being an ovarian tumour (i.e., an epithelial ovarian tumour selected from serous cystadenoma, borderline serous tumour, serous cystadenocarcinoma, mucinous cystadenoma, borderline mucinous tumour, mucinous cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma, clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner tumour. The ovarian tumour marker genes of the invention were identified using SAGE (serial analysis of gene expression) and were found to be overexpressed in a broad variety of ovarian epithelial tumour cells relative to normal ovarian epithelial cells. The marker genes are implicated in immune response pathways, in the regulation of cell proliferation and in protein folding, and many of these are membrane-localised or secreted. In addition to their use as diagnostic and prognostic markers, the ovarian tumour marker genes or their encoded proteins may be used as therapeutic targets for the treatment and prevention of ovarian cancer. (Sequences ABA83081-ABA83122, ABA83180, ABA83182 and ABA83184 represent the ovarian tumour marker genes of the invention.

XX Sequence 1147 BP: 210 A; 365 C; 425 G; 147 T; 0 other;

Alignment Scores:
Pred. No.: 8.58e-111 Length: 1147
Score: 1581.00 Matches: 316
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.68% Mismatches: 0
Query Match: 99.75% Indels: 0
DB: Gaps: 0

US-09-827-854-19 (1-317) x ABA83113 (1-1147)

QY 1 MettysValIleuTTPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 46 ATGAAAGTTCTGTGGGCTGCTGTCATTCCTCGCAGATGCCAAGCAAGTG 105
QY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnThrGlnSer 40
DB 106 GAGCAAGCGGTGGAGACAGACCCGAGCCGAGCCGACGAGACCGAGTGGCAGAGC 165
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyAArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 166 GGCAGCGCTGGAACTGGCACTGGCTTTGGGATTCCTGGCGTGGGTGCAGACA 225
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 226 CTGTCTGAGCGAGTGCAGAGAGAGCTGCTCAGTCCACAGTCACCCAGAACTAGAGCGC 285
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluLeu 100
DB 286 CTGATGACGAGACCATGAAGGAGTTGAAGCCCTACAAATCGAAGTGGAGAACAACTG 345
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 346 ACCCGGCGGCGAGAGACCGGCGCTGTCAAAGAGCTGCAGCGCGGCACAGGCC 405
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 406 CGGCTGGGCGGAGCATGAGAGAGCTGTGCGCGCTGTGGTACGATACCGCGGAGAGTG 465
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValAlaArgLeuAlaSerHisLeuArg 160
DB 466 CAGGCCATGCTCGGCGCAGAGACCGAGAGAGTGGGGTGGCCCTCCCACTGCGCC 525
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180

|||||:|||||
DB 526 AAGCTGGTAGCGGCTCTCCGATGCCATGACCTGCAGAACCCCTGGAGTGTAC 585
QY 181 GlnAlaGlyAlaArgGluGluAlaGlnArgLysLeuSerAlaIleArgLysValGlu 200
DB 586 CAGCGCGGCGCCGAGAGGCGCGAGCGGCTCAAGCGCATCCGAGACCGCTGGG 645
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
DB 646 CCCCTGGTGGAAACAGGCGCGGTGCGGCCCCACATGTGGGCTCCCTGGCGGCACGG 705
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
DB 706 CTACAGGAGCGGCGCCAGGCTGGGCGAGCGCTGCGCGCGCGCATGAGAGATGGGC 765
QY 241 SerArgTrpArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 766 AGCGGACCCCGACCGCTGGAGAGGTGAAGAGCAGGTGGGAGGTGGCGCCAG 825
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
DB 826 CTGAGAGAGAGCGCCGACAGATACCTGCAAGCCCAAGCCCTTCACAGCCCGCTCAG 885
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 886 AGCTGTTGCAAGCCCTGTGTGAAGACATGCAGCGCCAGTGGCGCGGCTGTGGAGAG 945
QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 946 GTGCAGGCTGCCGTGGGACACAGCGCCCTGTGGCCAGCAATTCAC 996

RESULT 5
AAF84315
ID AAF84315 standard; cDNA; 1156 BP.
XX
AC AAF84315;
XX

DT 21-JUN-2001 (first entry)

XX Human ApoE3 coding sequence.

XX Human; ApoE3; Alzheimer's disease; arteriosclerosis; ss.

OS Homo sapiens.

FH Key Location/Qualifiers
FT CDS 61..1014
FT /*tag= a
FT /product= "Human ApoE3"

PN JP2001017028-A.

XX 23-JAN-2001.

PD 28-APR-2000; 2000JP-0128919.

PF 06-MAY-1999; 99JP-0125647.

XX (MITU) MITSUBISHI CHEM CORP.

XX WPI; 2001-285406/30.

DR P-PSDB; AAB80997.

PT New apoe humanized mammalian cell useful for screening for agents
PT useful for treating or preventing Alzheimer's disease and
PT arteriosclerosis -

PS Disclosure; Page 13-14; 22pp; Japanese.

CC The present invention relates to an ApoE humanised mammalian cell. The
CC present sequence is the coding sequence for human ApoE3, which was used
CC in the method of the present invention. The ApoE humanised mammalian cell
CC can be used for screening for agents useful for treating or preventing

CC Alzheimer's disease and arteriosclerosis.

XX Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other;

Alignment Scores:

Pred. No.:	8.65e-111	Length:	1156
Score:	1581.00	Matches:	316
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.68%	Mismatches:	0
Query Match:	99.75%	Indels:	0
DB:	22	Gaps:	0

US-09-827-854-19 (1-317) x AAF84315 (1-1156)

```

QY 1 MetlyValLeuTPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIleVal 20
Db 61 ATGAAGCTTCTGTGGCGTGGCTGTGCTCAATTCCTGGCAGATGCGCAGCCAAAGTGT 120
QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAGAGCGGTGAGACAGAGCCGAGCCGAGCTGCCAGACAGACAGAGTGGCAGAGC 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlnArgPheTrpAspTyrIleuArgTrpAlaGlnThr 60
Db 181 GGCAGCGCTGGAGACTGGCTGGCTGTGGGATTAACCTGGCTGGGATGAGACA 240
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGCTGTGACAGGTGACAGAGAGAGCTGCTCAAGTCCAGAGTCCAGAGACTGAGAGCG 300
QY 81 LeuMetAspGluThrMetGlyGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu 100
Db 301 CTGATGAGCAGAGACCATGAGAGAGTTGAAGGCTTCAAAATGCGAACTGAGAGAACACTG 360
QY 101 ThrProValAlaGluGlnGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 361 ACCCGGTGGGAGAGACGAGGAGCGGCTGTCCAAGAGAGCTGACAGCGGCGAGGCC 420
QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlnArgLeuValGlnTyrArgGlyGluVal 140
Db 421 CGGCTGGCGCGGACATGAGAGAGCTGTGGCGGCTGTGGTGAATACCGGCGGAGAGTGT 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTGCGCCAGAGCAGCAGAGAGCTGCGGCTGCCCTCCACCTCCGCGC 540
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValLys 180
Db 541 AAGCTGCTAAGCGGCTCTCCGCGATGCCGATGACCTGCAAGAGCGCTGGCAGGTGAC 600
QY 181 GlnAlaGlnAlaArgGluGlnAlaGlnArgGlyLeuSerAlaIleArgGlnArgGlnGly 200
Db 601 CAGCGCGGCGCGCGAGGCGCGGCGGCGCTCACGCGCATGCCGAGCGCGCTGGGG 660
QY 201 ProLeuValGluGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGTGGGAAACAGGCGCGGCTGGGCGGCCCACTGTGGGCTCTCTGGCGGACAGCG 720
QY 221 LeuGlnGlnArgAlaGlnAlaThrGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
Db 721 CTACAGGAGCGGCGCGAGGCTGGGCGAGCGGCGGCGGCGGAGTGGAGAGATGGGCG 780
QY 241 SerArgTrpArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaLys 260
Db 781 AGCCGGAGCCCGCAGCCGCTGGAGAGGTGAAGAGAGAGTGGGAGAGTGGCGCCCAAG 840
QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 841 CTGAGAGGAGCGAGCGCCAGCATATGAGCTGAGCGCGAGGCTTCCAGAGCCCGCTCAAG 900
QY 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
Db 901 AGCTGTTCGAGCCCTGTGGTGAAGACATGAGCGCCAGTGGGCGGCGGCTGTGTGAGAAAG 960

```

QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGAGAGCTGCGGTGGCGACAGCGCCGCTGTGGCCAGAGCAATATCAC 1011

RESULT 6

AAD22048

AAD22048

12-FEB-2002 (first entry)

Human apolipoprotein E (apoE) isoprotein, apoE3 DNA.

Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;

hypertriglyceridaemia; low density lipoprotein; LDL; ds.

Homo sapiens.

Key Location/Qualifiers

sig_peptide /product= "Human apoE isoprotein, apoE3"

mat_peptide /tag= b

/product= "Mature human apoE isoprotein, apoE3"

W0200177136-A1.

18-OCT-2001.

06-APR-2001; 2001WO-US11358.

06-APR-2000; 2000US-0544386.

04-OCT-2000; 2000US-0679088.

05-APR-2001; 2001US-0827854.

(KOSP-) KOS PHARM INC.

(UTBO-) UNIV BOSTON.

Zannis VI, Kypros KE;

WPI: 2002-010885/01.

P-PSDB: AAE13294.

New apolipoprotein E polypeptide and nucleic acid, useful for lowering

cholesterol, delaying the onset of or treating atherosclerosis in

mammal, without inducing hypertriglyceridaemia

Claim 14: Page 81; 91pp; English.

The present sequence is a human apolipoprotein E (apoE)

isoprotein, apoE3 DNA. The apoE lipoproteins are useful for

lowering cholesterol, delaying the onset of atherosclerosis,

treating or regressing atherosclerosis without inducing

hypertriglyceridaemia, in a mammal lacking an endogenous,

normally functioning apoE gene or low density lipoprotein (LDL)

receptor or is at risk for developing atherosclerosis due to

accumulation of lipoprotein remnants in the bloodstream or having

a defect in remnant removal.

Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other;

Alignment Scores:

Pred. No.:	8.65e-111	Length:	1156
Score:	1581.00	Matches:	316
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.68%	Mismatches:	0
Query Match:	99.75%	Indels:	0
DB:	24	Gaps:	0

QY 121 ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 421 CGCTGGGCGGACATGAGAGAGCTGTGGCGCGCTGTGGCTAGTACCGCGGAGAGTG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCCATGCTCGGCCAGACGACGAGAGAGTGGGGTGGCCCTCCGCTCCACCTGCGC 540
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 541 AAGCTGTGAAGCGGCTCTCCCGATGCCGATGACCTGCAGAGATGCTCTCCAGTGTAC 600
QY 181 GluAlaGlyAlaArgGlnGluGluArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
DB 601 CAGGCCGCGGCGCGGAGGCGCGGAGCGGCGCTCAGCGCATCCGCGAGCGCTGGGG 660
QY 201 ProLeuValGlnGlnGlnGlnValArgValAlaAlaThrValGlySerLeuAlaGlyPro 220
DB 661 CCCTGTGGACAGCGGCGCGCTGCGGCGCGCTGTGGCTCTCCGCGGCGCAGCG 720
QY 221 LeuGlnGluArgAlaGlnAlaTTPGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
DB 721 CTACAGAGAGCGGCGCGAGCGCTGGGCGAGCGGCTGCGCGCGGATGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValAlaLys 260
DB 781 AGCGGAGCGCGCGACCGCTGAGAGAGGTAGAGAGCGGCGGAGGTGCGCGCCAG 840
QY 261 LeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
DB 841 CTGGAGAGAGAGCGCGCGCGATAGCGCTGAGAGCGCGCTTCCAGCGCGCTTAAG 900
QY 281 SerTrpHeGluProLeuValGluAspMetGlnArgGlnTrrPalaglyLeuValGluLys 300
DB 901 AGCTGGTTGACGCCCTCGTGGAAGACATGACAGCGCGCGGCGGCTGTGGAGAG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 961 GTGCAGGCTGCGGTGGCACGAGCGCGCGCTGTGCCAGCGACATTCAC 1011
RESULT 8
AAD22049
ID AAD22049 standard; DNA: 1156 BP.
AC AAD22049;
XX
XX 12-FEB-2002 (first entry)
DT
XX
DE Human apolipoprotein E (apoE) isoprotein, apoE2 DNA.
XX
XX Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;
KW hypertriglyceridaemia; low density lipoprotein; LDL; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 61..1014
FT FT /tag= a
FT FT /product= "Human apoE isoprotein, apoE2"
FT FT sig_peptide 61..114
FT FT /tag= b
FT FT mat_peptide 115..1011
FT FT /tag= c
FT FT /product= "Mature human apoE isoprotein, apoE2"
XX
PN WO200177136-A1.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-US11358.
XX
PR 06-APR-2000; 2000US-0544386.

PR 04-OCT-2000; 2000US-0679088.
PR 05-APR-2001; 2001US-0827854.
XX
XX (KOSP-) KOS PHARM INC.
PA (UYBO-) UNIV BOSTON.
XX
XX Zannis VI, Kyreos KE;
XX WPI; 2002-010885/01.
DR P-PSDB; AAEI3295.
PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT cholesterol, delaying the onset of or treating atherosclerosis in
PT mammal, without inducing hypertriglyceridaemia -
XX
XX Claim 14; Page 81-82; 91pp; English.
XX
CC The present sequence is a human apolipoprotein E (apoE)
CC isoprotein, apoE2 DNA. The apoE lipoproteins are useful for
CC lowering cholesterol, delaying the onset of atherosclerosis,
CC treating or regressing atherosclerosis without inducing
CC hypertriglyceridaemia, in a mammal lacking an endogenous,
CC normally functioning apoE gene or low density lipoprotein (LDL)
CC receptor or is at risk for developing atherosclerosis due to
CC accumulation of lipoprotein remnants in the bloodstream or having
CC a defect in remnant removal.
XX
SQ Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;
XX
Alignment Scores:
Pred. No.: 3,48e-110 Length: 1156
Score: 1573.00 Matches: 315
Percent Similarity: 99.68% Conservative: 1
Best Local Similarity: 99.37% Mismatches: 1
Query Match: 99.24% Indels: 0
DB: 24 Gaps: 0
US-09-827-854-19 (1-317) x AAD22049 (1-1156)
QY 1 MetLysValLeuTrrPalalalaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 61 ATGAGGTTCTGTGGGCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrrPginSer 40
DB 121 GAGCAACGGGTGAGAGAGAGCGCGAGCCCGAGCTGGCGCAGAGACCGAGTGGCAGGC 180
QY 41 GlyLnaArgTrrPgluLeuAlaLeuGlyArgPheTrrAspTrrLeuArgTrrPValGlnThr 60
DB 181 GGCAGGCGCTGGAACTGGCACTGGCTTTTGGGATTAACCTGGCTGGTGGCAGACA 240
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 241 CTGTGTGAGCAGGTGAGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrrLysSerLysLeuGlnGlnLeu 100
DB 301 CTGATGACGACGACCATGAAAGGAGTTGAAAGCCCTACAAATCGGAACCTGGAGAACTG 360
QY 101 ThrProValAlaGlnGluTrrArgAlaArgLeuSerLysGlnGlnGlnAlaAlaGlnAla 120
DB 361 ACCCGGTGGGAGAGAGAGCGGCGAGCTGTCCAAAGAGCTGCAGCGCGCGCAGGCC 420
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 421 CGGCTGGGCGCGGACATGAGAGAGCTGTGGCGCGCTGTGGCTAGTACCGCGGAGAGTG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCCATGCTCGGCCAGACGACGAGAGAGTGGGGTGGCCCTCCGCTCCACCTGCGC 540
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180

OY 301 ValGlnAlaIaValaGlyThrSerAlaAlaProValProSerAspAsnHis 317
|||||
Db 962 GTGACAGGTGGCGTGGGACACGCGCCGCTGTGCCAGGACATCACC 1012

RESULT 11
ABN95746
ID ABN95746 standard; DNA; 1157 BP.
XX
AC ABN95746;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #2244 used to diagnose liver cancer.
XX
KW Gene: liver cancer; ds: hepatocellular carcinoma; hepatotropic;
metastatic liver tumour; cytostatic; expression profile; disease state;
disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30589.
XX
PR 02-OCT-2000; 2000US-237054P.
XX
PA (GENE-) GENE LOGIC INC.
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
DR WPI, 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer,
hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -
XX
PS Claim 1; SEQ ID NO 2244; 298bp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytosolic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SO Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;

Alignment Scores:
Pred. No.: 4,15e-110 Length: 1157
Score: 1572.00 Matches: 314
Percent Similarity: 99.37% Conservative: 1
Best Local Similarity: 99.05% Mismatches: 2
Query Match: 99.18% Indels: 0
Gaps: 0
DB: 24

US-09-827-854-19 (1-317) x ABN95746 (1-1157)

OY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
Db 62 ATGAAGGTTCTGTGGCGTGGCTGTGCATTCTCGGCGAGATGCGACGCCAAGGTG 121

OY 21 GlnGlnAlaValaGlyThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
|||||
Db 122 GAGCAAGCGGTGGAGACAGAGCGGAGCGGAGCTGGCGCACAGACGAGGGGAGAC 181

OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
Db 182 GGCCAGCGCTGGAACTGTGCACCTGGTCTTTGGGATTACCTCGCTGGGTGCAGACA 241

OY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
Db 242 CTGTGTGACAGGTGCAGAGGAGGCTGCTACCTCCCAAGTACCACCAAGACTGAGGGG 301

OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
|||||
Db 302 CTGATGAGCGAACCATTGAGAGGTTGAGGCTTACAAATCGGAAGTGGAGGACACTG 361

OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
|||||
Db 362 ACCCGGTAGCGAGAGAGACGCGGCGACGCTGTCCAAAGAGCTGCAGACGCGCAGGCC 421

OY 121 ArgLeuGlyAlaAspMetGluLysValCysGlyArgLeuValGlnTrpArgGlyValAl 140
|||||
Db 422 CGGCTGGGCGCGACATGAGGAGCTGTGGCGCGCTGTGTGACGTACCGCGGAGGTG 481

OY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgValAlaSerHisLeuArg 160
|||||
Db 482 CAGGCCATCTCTGGCGACAGACCGAGAGCTGCGGGTCTGCGCTCCCACTCGCGCC 541

OY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
Db 542 AAGCTGCGTAAGCGGCTCTCCGCGATCCCGATGACCTCCAGAAAGCGCTGCGACGTAC 601

OY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaHisLeuArgLeuGly 200
|||||
Db 602 CAGGCGGGGCGCGGAGGCGCGCGAGCGCGCTCAGCGCATCCGACAGGCGCTGGGGG 661

OY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyLysPro 220
|||||
Db 662 CCCCTGTGGAACAGAGCGCGCTGGGCGGCGGCTGTGGCTCTGCGCGGCGGCGGCGG 721

OY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluLysMetGly 240
|||||
Db 722 CTACAGAGCGGCGCGAGGCTGTGGCGAGCGGTGGCGCGCGATGGAGAGATGGGCG 781

OY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
|||||
Db 782 ACTCGGACCGCGAGCGCTGTGACAGAGGAGGAGCAAGTGGCGGAGCTGGCGCGCAAG 841

OY 261 LeuGlnGluGlnAlaGlnGlnThrLeuArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
|||||
Db 842 CTGGAGAGAGCGCGCGAGCATACGCTGCGAGGCGGAGGCTTCCAGGCGCGCTCAAG 901

OY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGlyLys 300
|||||
Db 902 AGCTGTTGACGCCCTGTGTGAACAATGACAGCGCGAGTGGCGGCGGTGTGTGAGAAAG 961

OY 301 ValGlnAlaIaValaGlyThrSerAlaAlaProValProSerAspAsnHis 317
|||||
Db 962 GTGACAGGTGGCGTGGGACACGCGCCGCTGTGCCAGGACATCACC 1012

RESULT 12
ABK64514
ID ABK64514 standard; DNA; 1157 BP.
XX
AC ABK64514;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human benign prostatic hyperplasia gene #409.
XX
KW Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
XX
OS Homo sapiens.

XX WO200212440-A2.
 PN 14-FEB-2002.
 XX
 XX 07-AUG-2001; 2001WO-US24708.
 PF
 XX 07-AUG-2000; 2000US-22323P.
 PR 05-JUN-2001; 2001US-0873319.
 XX
 XX (GENE-) GENE LOGIC INC.
 PA (NISB) JAPAN TOBACCO INC.
 XX
 PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
 DR WPI: 2002-257476/30.
 XX
 XX
 PT Identifying drugs for and diagnosing benign prostatic hyperplasia, by
 PT detecting expression levels of one or more genes in prostate cells from
 PT patient that are differentially regulated compared to normal prostate
 PT cells -
 XX
 XX
 PS Disclosure: Page 239-240; 444pp; English.
 XX
 CC The invention relates to a method of diagnosing (I) the onset or
 CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
 CC or identifying an agent that modulates the onset or progression of BPH.
 CC The method is based on changes in gene expression in BPH tissue isolated
 CC from patients exhibiting different clinical states of prostate
 CC hyperplasia as compared to normal prostate tissue. (I) comprises
 CC detecting the expression levels of one or more genes in prostate cells
 CC from the subject that are differentially regulated compared to normal
 CC prostate cells. (II) comprises preparing a first gene expression profile
 CC of BPH cells or BPH-like cell population, exposing the cells to the
 CC agent, preparing a second gene expression profile of the agent exposed
 CC cells, and comparing the first and second gene expression profiles.
 CC (I) is useful for diagnosing the onset or progression of BPH. (II) is
 CC useful for identifying an agent that modulates the onset or progression
 CC of BPH. The methods are useful to present information identifying
 CC the expression level in a tissue or cells, by comparing the expression
 CC level of genes given in the specification in the tissue or cells to the
 CC level of expression of gene in the database, and displaying the
 CC expression levels of at least one gene in the tissue or cell sample
 CC compared to the expression level in BPH. Agents using (II) are useful for
 CC treating BPH or prostate cancer. ABK64106-ABK64860 represent human
 CC benign prostatic hyperplasia gene sequences of the invention.
 XX
 SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;

Alignment Scores:
 Pred. No.: 4,15e-110 Length: 1157
 Score: 1572.00 Matches: 314
 Percent Similarity: 99.37% Conservative: 1
 Best Local Similarity: 99.05% Mismatches: 2
 Query Match: 99.18% Indels: 0
 DB: 24 Gaps: 0

US-09-827-854-19 (1-317) x ABK64514 (1-1157)

OY 1 MetLysValLeuThrPAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 62 ATGAAAGTTCTGGGCTCGTGGTGCATTCCTGCGAGGATCCAGGCCAAGGTTG 121
 OY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 122 GAGCAAGCGGTGAGAGCAAGCCGAGCCGACCTCGCCAGCAGACCAAGTGGCAGAGC 181
 OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB 182 GGCAGAGCGCTGGAACTGCGACTGGGTCTTTGGATTACCTGCGGTGGGAGACA 241
 OY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80

DB 242 CTGTCTGACAGGTGACGAGAGAGCTGCTACGCTCCCAAGTCAACCCAGAACTGAGGGCG 301
 OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 DB 302 CTGATGAGAGACCATGAAAGAGGTGAAGGCCCTACAATAATGCGAACTGGAGAACAACTG 361
 OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 362 ACCCCGCTAGCGAGAGAGACCGGGGACGCGCTGTCCAAAGAGCTGACAGCGGCGAGGCC 421
 OY 121 ArgLeuGluValAlaAspMetGluAspValLysGlyArgLeuValGlnTrpArgGlyGluVal 140
 DB 422 CGGCTGGCGCGGACATGAGAGACGTGTGCGCGCTGTCAGTACCGCGCGAGGTG 481
 OY 141 GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 482 CAGGCATGCTGCGGACAGAGACCCAGAGAGCTGGGGTGGCTGCTCCACCTGGCCG 541
 OY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaTyr 180
 DB 542 AAGCTCGTAAAGCGGCTCTCCGCGATCCCGATGACCTGCAGAAAGCGCTGGCAGTGTAC 601
 OY 181 GlnAlaGluValAlaArgGluGluValAlaGluArgGlyLeuSerAlaTrpLeuArgLysLeuGly 200
 DB 602 CAGGCGGGGCGCGGAGGCGCGGAGCGCGGCTCAGCGCATCCGCGAGCGCTGGGG 661
 OY 201 ProLeuValGluGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
 DB 662 CCCCTGTGGAACAGGCGCGCGTGGCGGCGCCGACATGTGGGCTCCCTGGCGGCGAGCGG 721
 OY 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluGly 240
 DB 722 CTACAGAGACGGGCGCGAGGCTGGGGCGAGCGGCTGCGCGCGGTGAGAGAGAGGGCG 781
 OY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 DB 782 AGTCGACCCCGGACCGCGCTGAGAGAGGTGAAGAGACAGTGGCGAGGTGGCGCGCAAG 841
 DB 842 CTGAGAGACAGGCGCCAGCAGATACCGCTGAGCGGAGCGCTTCCAGCGCGCGCTCAAG 901
 OY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 DB 902 AGCTGTTGGAGCCCTGTGTGAAGACATGCAAGCCGAGTGGCGCGGTGGTGGAGAG 961
 OY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 962 GTGCAGGCTGCGGTGGGCAACAGCGCGCCCTGTGCCACGACACATCAC 1012

RESULT 13
 ABL65450
 ID ABL65450 standard; DNA; 1157 BP.
 XX
 AC ABL65450;
 XX
 XX 15-MAY-2002 (first entry)
 DT
 XX
 XX Lung cancer related gene sequence SEQ ID NO:3787.
 DE
 XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumor; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US10838.
 XX

PR 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 29-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 PA (AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Sopet DR, Weaver Z;
 XX
 DR WPI; 2002-188264/24.
 XX
 XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 XX Claim 1; SEQ ID 3787; 44pp; English.
 PS
 XX
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate and pancreatic cancer.
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilm's tumour.

XX
 SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other:
 Alignment Scores:
 Pred. No.: 4.15e-110 Length: 1157
 Score: 1572.00 Matches: 314
 Percent Similarity: 99.37% Conservative: 1
 Best Local Similarity: 99.05% Mismatches: 2
 Query Match: 99.18% Indels: 0
 DB: 24 Gaps: 0
 US-09-827-854-19 (1-317) x ABL65450 (1-1157)
 QY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 62 ATGAAGGTTCTGTGGCGTGGCTGTGGTGCACATTCCTGCGAGATGCCAGGCCAAGG 121
 QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluPheGlnSer 40
 DB 122 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCGCCAGACAGACGAGTGCAGACA 181
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB 182 GCCAGCGCTGGAACTGGCACTGGGTGCTTTGGATTACTGCTGCGGTGCAGACA 241
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB 242 CTGTGTGACAGAGTCCAGAGAGAGCTGCTCAGCTCCCAAGTACCAGACTGAGAGCGC 301
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
 DB 302 CTGATGAGACAGACCATGAGAGAGTTGAAGGCCCTCAATCGAATCGAAGTGGAGAACACTG 361
 QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 362 ACCCGGTAGCGAGAGAACCGCGGCGACGGCTGCTCAAGAGACTGCAGACGCGCAGAGCC 421
 QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnThrArgGlyGluVal 140
 DB 422 CGGCTGGGCGGACATGAGAGAGCTGTGGCGGCTGTGTGCAATACCGGCGGCAAGGTG 481
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 482 CAGGCCATGCTGGCCGACAGAGCACCGAGAGCTGGCGGTGCGGCTGCCACCTGGCGC 541
 QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 DB 542 AAGCTGCGTAAGCGGCTCTCCGCGATCCCGATGACCTGCAGAACGCGCTGCACTGTAC 601
 QY 181 GlnAlaGlyAlaArgGluGluValAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 602 CAGCGCGGCGGCGGAGGCGGCGGCGGCGGCTCAGCGCCATCCGAGAGCGCTGGGG 661
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 DB 662 CCCCTGTGTGAACAGAGCGCGGCTGGCGGCGGCTGAGTGGCTCCCTGCGCGGCGAGCCG 721
 QY 221 LeuGlnGluArgAlaGlnAlaThrProGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
 DB 722 CTACAGGAGCGGCGGCGGAGGCTGGGGCGAGCGCGCGCGCGGATGAGAGATGGGCG 781
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAlaArgAlaLys 260
 DB 782 AGTCGAGACCGCGGCGGCTGGAGAGTGAAGAGAGAGTGGCGGAGTGGCGGCGCAAG 841
 QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 DB 842 CTGAGAGAGAGCGGCGGACAGATAGCTGCGAGCGCGGAGGCTTCAGAGCGCGCTCAAG 901
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 DB 902 AGCTGTTTCGAGCGGCTGTGTGGAACATGACGCGCCAGTGGCGGCTGTGTGGAAG 961

Oy	301	VaIgfnaIaIaVaIcIyThSerAlaIaIaProValProSerAspAsnHis	317
Db	962	GTGcAGcGTGCGTGGcCACcAGcCGcCGcCGCTGTGcCCAGcGACATTCAC	1012
RESULT 14			
ID	AANS0450		
XX	AANS0450 standard; DNA; 1110 BP.		
AC	AANS0450;		
XX			
DT	09-JAN-1992 (first entry)		
XX			
DE	Sequence encoding human apolipoprotein E.		
XX			
KW	Hyperlipaemia; ds.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	15..968	
FT		/*tag= a	
FT	mat_peptide	69..965	
FT		/*tag= b	
XX			
PN	JP60118189-A.		
XX			
PD	25-JUN-1985.		
XX			
PE	29-NOV-1983; 83JP-0224980.		
XX			
PR	29-NOV-1983; 83JP-0224980.		
XX			
PA	(MITU) MITSUBISHI CHEM IND KR.		
XX			
DR	WPI; 1985-188003/31.		
DR	P-PSDB; AAP51204.		
XX			
PT	DNA fragment - conty. DNA which codes human apolipoprotein E for		
PT	treatment of hyperlipaemia.		
XX			
PS	Claim 3; Page 484; 8pp; Japanese.		
XX			
CC	The sequence may be used to produce the apolipoprotein E, useful in the		
CC	treatment of hyperlipaemia.		
XX			
SO	Sequence 1110 BP; 198 A; 355 C; 414 G; 143 T; 0 other;		
Alignment Scores:			
	Pred. No.:	4, 71e-110	Length: 1110
	Score:	1571.00	Matches: 314
	Percent Similarity:	99.37%	Conservative: 1
	Best Local Similarity:	99.05%	Mismatches: 2
	Query Match:	99.12%	Indels: 0
	DB:	6	Gaps: 0
US-09-827-854-19 (1-317) x AANS0450 (1-1110)			
OY	1	MetIysValIleuTrpAlaIaIaIeuleuValThrPheIeuAlaGlyCysGlnAlaIaIysVal	20
Db	15	ATGAAGGTTCTGTGGCGCTGCTGCTGTGCATTCCTGGCAGATGCCAGGCCAAGGTG	74
OY	21	GlUGlnAlaValaIGluThrGluProGluIuProGluIeuAurGlnGlnInThrGluTrpGlnSer	40
Db	75	GAGCAAGCGGTGGAGACAGAGCGCGGAGCCGACGCTGGCCGACAGACCGAGTGGCAGAGC	134
OY	41	GlyGlnaIaGTrPGluIeuAlaIeuleuGlyIaTrpPheTrpAspTyIeUaArgTrpValGlnThr	60
Db	135	GGCCAGCCCTGGGAACCTGGCAGCTGGTGCTTTTGGATTCACCTGCCCTGGGTGCAGACA	194
OY	61	LeuSerGluGlnValaGlnGluGluIeuleuSerSerGlnValaTrhGlnGluIeuAurGla	80
Db	195	CTGTCTGAGCAGCATGTCAGAGAGAGCTGCTACGCTCCAGGTCAACCCAGAACTGAGGCG	254

QY	81	LeuueTAspGIuTrmMeLysGIuLeuYAlaIarTyrLysSerGIuLeuGIuGIuLeu	100
Db	225	CTGATGGACGAGACCATGTAAGAGATTGAAGCCCTACAAATTCGGAACCTGGAGGAACAAC	314
QY	101	ThrProValAlaGIuGIuTrhArgAlaArgLeuSerLysGIuLeuGIuAlaAlaGIuAla	120
Db	315	ACCCGGTGGGGAGAGAGACGCGGGCCAGGCTGTCCAAAGAGAGCTGCAGGCGCCACAGGCC	374
QY	121	ArgLeuGIuAlaAspMetGIuAspValCysGIuArgLeuValGIuTyrArgGIuGIuVal	140
Db	375	CGGCTGGGGCGGACATGAGAGAGAGTGTGGGGCCGCTGTGTGCAGTACCCGGCGAGGTG	434
QY	141	GIuAlaMetLeuGIuYGIuSerThGIuGIuLeuGIuArgAlaArgLeuAlaSerHisLeuArg	160
Db	435	CAGGCATAGCTGTGGCCAGAGACACCGAGAGAGTGGGGGTGGCCCTCGCCTCCACCTGGCC	494
QY	161	LysLeuArgGIuAlaArgLeuLeuArgAspAlaAspAspLeuGIuLysArgLeuAlaValTyr	180
Db	495	AAGCTGGCTAAGCGGCTCTCTCGCGAGTCCGATGACCTGCAGAAAGCCCTGGCAGTGTAC	554
QY	181	GIuAlaGIuAlaIaArgGIuGIuAlaGIuArgGIuLeuSerAlaIleArgGIuArgLeuGIu	200
Db	555	CAGCGCCGGGGCCCGAGAGGGCGGACCGGGCCCTCAGCCCAATCCCGCAGCGCCCTGGGG	614
QY	201	ProLeuValGIuGIuGIuArgYValArgAlaAlaIaIaIaIaIaIaIaIaIaIaGIuLeuGIuPro	220
Db	615	CCCCGTGGTGACAGAGGGCGGTGGGGCCGCCCACTTGGGCTCTCTCGCCGGCCAGCCG	674
QY	221	LeuGIuGIuAlaArgAlaGIuAlaIaTrpGIuGIuArgLeuArgAlaArgMetGIuGIuMetGIu	240
Db	675	CTACAGAGAGCGGGCCAGAGCCCTGGGGGAGAGGGCTGGCGCGCGGAGTGAAGAGATGGGC	734
QY	241	SerArgTrhArgAspArgLeuAspGIuValLysGIuGIuValAlaGIuValaIaGIuValaIaLys	260
Db	735	AGCCGAGCCCGCCGACCCCTGGAGAGAGTGAAGAGCAGCGTGGCGGAGGTGGCGGCCAG	794
QY	261	LeuGIuGIuGIuAlaGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu	280
Db	795	CTGAGAGGAGCGAGCGCCAGCAAGATCGCTCTCAGGCCGAGGCCCTTCCAGGCGCCCTCAG	854
QY	281	SerTrpPheGIuProLeuValGIuAspMetGIuArgGIuTrpAlaGIuLeuValGIuLys	300
Db	855	AGCTGGTTCGAGCCCTGTGTGGAAAGACATGCAGCGCCAGTGGGGCGGTGTGGAGAG	914
QY	301	ValGIuAlaIaValGIuTrhSerAlaAlaProValProSerAspAsnHis	317
Db	915	GTCGAGGCTGCGTGGGACACAGCGCGCCCTGTGTGCCAGCAATATC	965
RESULT 15			
AAAF84316			
ID AAF84316 standard; cDNA; 1156 BP.			
XX	AAAF84316:		
XX	AC		
XX	21-JUN-2001 (first entry)		
XX	DE		
XX	Human Apoe4 coding sequence.		
XX	Human; Apoe4; Alzheimer's disease; arteriosclerosis; ss.		
XX	Homo sapiens.		
XX	Key	Location/Qualifiers	
XX	FT CDS	61..1014	
XX	FT	/+tag= a	
XX	FT	/product= "Human Apoe4"	
XX	JP2001017028-A.		
XX	23-JAN-2001.		
XX	28-APR-2000; 2000JP-0128919.		

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:17:52 ; Search time 34.7343 Seconds
(without alignments)
2798.866 Million cell updates/sec

Title: US-09-827-854-19

Perfect score: 1885

Sequence: 1 MKVLMALIVFLAGCAKV.....VEKVOAVGTSAPVPSDNH 317

Scoring table: BIOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 862724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+P2n.model -DEV=XLP
-O=/cgn2_1/USPFO.spool/US09827854/runat_11032003_101611_27510/app_query.fasta.1.3576
-DB-issued.patents.NA -QEMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdl
-LIST=45 -DOCCALIGN=200 -THR SCORE=0.0 -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MITLEN=0 -MAXLEN=2000000000
-USER=US09827854.CGN.1.1.2254 &runat_11032003_101611_27510 -NCPU=6 -ICPU=3
-NO.XLPPY -NO.MMAP -LARGEOUERT -NEG-SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCUTS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles.seq:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1572	99.2	1157	1	US-07-709-949-1
2	1134	71.5	1126	4	US-08-949-155-5
3	1134	71.5	1126	4	US-09-819-964-5
4	998	63.0	4267	4	US-08-949-155-51
5	998	63.0	4267	4	US-09-819-964-51
6	980	61.8	660	2	US-08-726-306A-28
7	507	32.0	330	1	US-07-849-389-6
8	370	23.3	252	3	US-08-617-256-24
9	370	23.3	252	4	US-09-287-141-24
10	370	23.3	252	4	US-09-431-613-24
11	370	23.3	252	4	US-09-504-245-24
12	370	23.3	252	4	US-09-287-682-24

13	370	23.3	252	4	US-09-287-679-24	Sequence 24, Appl
14	370	23.3	252	4	US-09-397-766-24	Sequence 24, Appl
15	370	23.3	252	4	US-09-287-681-24	Sequence 24, Appl
16	370	23.3	252	4	US-09-495-444-24	Sequence 24, Appl
17	176.5	11.1	842	4	US-08-952-796-1	Sequence 1, Appl
18	173.5	10.9	801	1	US-07-959-946-4	Sequence 4, Appl
19	173.5	10.9	801	1	US-08-333-577-4	Sequence 4, Appl
20	173.5	10.9	801	5	PCT-US92-08634-4	Sequence 4, Appl
21	159	10.0	5661	4	US-08-938-105-2	Sequence 2, Appl
22	156.5	9.9	964	1	US-08-448-606-5	Sequence 3, Appl
23	155	9.8	3256	2	US-08-968-751-3	Sequence 3, Appl
24	154	9.7	1879	4	US-09-750-580-2	Sequence 2, Appl
25	153.5	9.7	863	1	US-08-448-606-7	Sequence 2, Appl
26	149	9.4	6306	1	US-08-195-487-3	Sequence 3, Appl
27	149	9.4	6306	5	PCT-US93-06160-3	Sequence 3, Appl
28	146.5	9.2	603	4	US-08-952-726-14	Sequence 14, Appl
29	145	9.1	516	1	US-09-183-861-34	Sequence 34, Appl
30	145	9.1	516	4	US-09-022-765-34	Sequence 3, Appl
31	145	9.1	6306	1	US-08-466-390-3	Sequence 3, Appl
32	145	9.1	6306	1	US-08-470-950-3	Sequence 3, Appl
33	145	9.1	6306	1	US-08-467-781-3	Sequence 3, Appl
34	145	9.1	6306	2	US-08-483-924-3	Sequence 3, Appl
35	144	9.1	1771	2	US-08-533-669A-7	Sequence 7, Appl
36	144	9.1	1771	2	US-08-511-872-1	Sequence 1, Appl
37	144	9.1	1771	4	US-09-183-861-7	Sequence 7, Appl
38	144	9.1	1771	4	US-09-023-765-7	Sequence 7, Appl
39	144	9.1	1960	2	US-08-533-306A-1	Sequence 1, Appl
40	144	9.1	1960	2	US-08-742-923A-1	Sequence 1, Appl
41	144	9.1	8789	1	US-08-328-254-5	Sequence 5, Appl
42	144	9.1	10136	1	US-08-353-700-2	Sequence 2, Appl
43	144	9.1	10136	5	PCT-US95-16216-2	Sequence 2, Appl
44	144	9.1	13121	4	US-08-961-527-126	Sequence 126, App
45	143.5	9.1	11236	1	US-07-853-913-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-07-709-949-1
Sequence 1, Application US/07709949
Patent No. 5472858
GENERAL INFORMATION:
APPLICANT: Attie, Alan D
APPLICANT: Gretch, Daniel G
APPLICANT: Sturley, Stephen L
APPLICANT: Beckage, Nancy E
TITLE OF INVENTION: Production of Recombinant Proteins in
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: P.O. Box 2113
CITY: Madison
STATE: WI
COUNTRY: USA
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/709,949
FILING DATE: 19910604
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 9629691801
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 1:


```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 1157 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FRAGMENT TYPE: N-terminal
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 62..1015
:
: NAME/KEY: misc.signal
: LOCATION: 62..119
: PUBLICATION INFORMATION:
: AUTHORS: McLean, J W
: JOURNAL: Journal of Biological Chemistry
: VOLUME: 259
: PAGES: 6498-6504
: DATE: 1984
:
: RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 1157
:
: US-07-709-949-1
:
: Alignment Scores:
: Pred. No.: 1,45e-146 Length: 1157
: Score: 1572.00 Matches: 314
: Percent Similarity: 99.37% Conservative: 1
: Best Local Similarity: 99.05% Mismatches: 2
: Query Match: 99.18% Indels: 0
: DB: Gaps: 1
:
: US-09-827-854-19 (1-317) x US-07-709-949-1 (1-1157)
:
: QY 1 MetLysValLeuTrpAlaIaLeuLeuValTrpPheLeuAlaGlyCysGlnAlaLysVal 20
: DB 62 ATGAAGGTTCTGTGGCTGCTGTGTCACATTCCTGCGAGATGCCAGCGCAAGGTC 121
:
: QY 21 GUGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluProGln 40
: DB 122 GAGCAAGCGGTGAGACAGACCGGAGCCGAGCTCGCGACAGACCGAGTGCGAGAGC 181
:
: QY 41 GYGLnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
: DB 182 GGGCAGCGCTGGGAATCGGCACCTGCTTTGGGATTACTGGCGCTGGGCGAGACA 241
:
: QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
: DB 242 CTGTCTGAGCAGAGTGCAGAGAGCTGCTACGCTCCCAAGTCAACCAAGAACTGAGGCGC 301
:
: QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100
: DB 302 CTGATGACAGACCAATGAAGGATTGAAGGCTTACAAATCGGAATGGAGGACAACATG 361
:
: QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
: DB 362 ACCCGCGTAGCGGAGAGAGACGGCGGACGGCTGTCCAAGAGCTGCAGACGGCGAGGCC 421
:
: QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
: DB 422 CCGCTGGCGCGGACATGAGAGACGTGTGCGCGCTGCTGTCAGTACCGCGCGAGGTG 481
:
: QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
: DB 482 CAGGCCATCTCGGCGACAGACCGAGAGCTCGGGTGGCTCCCTCCACCTGCGCC 541
:
: QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
: DB 542 AAGCTCGTGAAGCGGCTCTCCCGGATCCCGATGACCTGCAGAGAGCGCTGGAGATGTC 601
:
: QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
```

```

: DB 602 CAGCGCGGCGCCGCGAGGCGCGCGGCTTCAGCGCCATCCGACGCGCTGGGG 661
:
: QY 201 ProLeuValGluGlnGlnLysArgValArgAlaIaIaThrValGlySerLeuAlaGlyGlnPro 220
: DB 662 CCCCTGGTGAAACAGGCGCGGCTGGCGGCGCCACTGTGGGTCTCCCTGGCGCGGCGCGG 721
:
: QY 221 LeuGlnAlaArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
: DB 722 CTACAGAGCGGCGCGGCGGCTGGCGGCGAGCGGCTGCGCGCGGATGAGAGATGGGCG 781
:
: QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
: DB 782 AGTCGAGACCGCGACCGCTGCGAGAGTGAAGAGACAGTGCAGGTGCGCGCGCAAG 841
:
: QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnIleArgLeuLys 280
: DB 842 CTGGAGAGCGAGCGCCAGCAGATACGCTGCAGGCGGACCTTCCAGGCGCGGCTCAAG 901
:
: QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
: DB 902 AGCTGGTTGAGGCGCTGGTGGAAGACATGACGCGCAGTGGCGCGGCTGGTGAGAGAG 961
:
: QY 301 ValGlnAlaIaValGlyThrSerAlaIaProValProSerAspAsnHis 317
: DB 962 GTGCAGGCTGCGCTGGGACACACGCGCGCCCTGTGCTCCAGGACACATCAC 1012
:
: RESULT 2
: US-08-949-155-5
: Sequence 5, Application US/08949155
: Patent No. 6271436
: GENERAL INFORMATION:
: APPLICANT: Piedrahlta, Jorge A
: TITLE OF INVENTION: Compositions and Methods for the
: NUMBER OF SEQUENCES: 51
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ARNOLD, WHITE AND DURKEE
: STREET: P.O. Box 4433
: CITY: Houston
: STATE: TX
: COUNTRY: US
: ZIP: 77210-4433
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/949,155
: FILING DATE: Concurrently Herewith
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/027,338
: FILING DATE: 11-OCT-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/046,094
: FILING DATE: 09-MAY-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Hidler, David W.
: REGISTRATION NUMBER: 41,071
: REFERENCE/DOCKET NUMBER: TAAK:177
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (512) 418-3000
: TELEFAX: (713) 789-2679
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1126 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: FEATURE:
```


NAME/KEY: CDS
LOCATION: 51..1001
US-08-949-155-5

Alignment Scores:

Pred. No.:	2.79e-103	Length:	1126
Score:	1134.00	Matches:	223
Percent Similarity:	83.75%	Conservative:	45
Best Local Similarity:	69.69%	Mismatches:	44
Query Match:	71.55%	Indels:	8
	4	Gaps:	3

US-09-827-854-19 (1-317) x US-08-949-155-5 (1-1126)

```

Qy 1 MetysValleuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 51 ATGAGGGTTCTGTGGTGTCTTGTGTGTACCTCTCCGACGATGCGGACAGAGAC 110
Qy 21 GluGlnAlaValGluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 111 GAGCCGGGG-----CCGCGCGCGGAGGTGCAAGTGTGTGGAGAGAGCCCAAG 158
Qy 38 TrpGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrp 57
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 159 TGGCAGGCGCACGCCCTGGGAGCAGGCCCTGGGCGCTTCTGGGATTACCTGGCTGG 218
Qy 58 ValGlnThrLeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGlu 77
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 219 GTGCAAGTCTCTGTCTGACCAAGTGCAGAGAGGTGCTGACACCAAGTCCACAGGAA 278
Qy 78 LeuArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyrlLysSerGluLeuGlu 97
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 279 CTGACGAGCTGATGATGAGAGAGAGCATGAAAGGTGAAGGCTTACCGGAGAGAGCTGAG 338
Qy 98 GluGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAla 117
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 339 GCGCAGCTGGGCGCCCTGACCCAGCAGAGCGCGCCCTCTCCAAAGAGCTGACGAGCG 398
Qy 118 AlaGlnAlaArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArg 137
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 399 GCGCAGGCGCGCGGTGGGCGCGGACATGAGACGTGCGCAACCGCTTGTCTCTACCGGC 458
Qy 138 GlyLeuValGlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSer 157
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 459 AGCGAGGTGCACACATGTGTGGCCAGACACCGAGGAGCTGCGGAGCGGCTGCTCC 518
Qy 158 HisLeuArgLysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeu 177
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 519 CACCTGCCCAAGCTGCCAGCGGCTGCTCCGCGACACCGAGAGACTGCAAGAGCGCTG 578
Qy 178 AlaValTyrglnAlaGlyAlaArgGluGlnGluAlaGluArgGlyLeuSerAlaIleArgGlu 197
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 579 GCGGTGTACAGCGGGGCTGCGGAGGGCGCGCGACGACGAGGAGGCGCTCCGCGAG 638
Qy 198 ArgLeuGlyProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAla 217
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 639 GCGCTCGGGGCGCGGTGTGTGAGCAGGCGCGGATTCGCGCGCGCTAGTACCAAGGGCC 698
Qy 218 GlyGlnProLeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGlu 237
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 699 GCGCAGCGCGGTGCGAGCGCGCGGAAAGCTGGGGCGCGAGAGCTGCGGAGCGCTGAG 758
Qy 238 GluMetGlySerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluVal 257
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 759 GAGATGGGACAGCGGACCGCGGACCGCTGATGATGATGATGATGATGATGATGATG 818
Qy 258 ArgAlaLysLeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 277
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 819 CGCACCAAAAGTGAAGAGACAGGCGCAGTTGGCGCTGCGAGCGCGAGGATTCACCGCC 878
Qy 278 ArgLeuLysSerTrpPheGlnProLeuValGluLysMetGlnArgGlnTrpAlaGlyLeu 297
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 879 CTCCTCAAAAGCTGTGTGAGCTCTCTGTGTGAGAGACATACGCGCGCAGTGGGCGGCTG 938
  
```

Qy 298 ValGluLysValGlnAlaValGlyLysValThrSerAlaAlaProValProSerAspAsn 316
 Db 939 GTGGAGAGAGATGACAGTGGCGCGTGGAGCATAGCTCTCCACCTGCGCGCCAGCATATAT 998

RESULT 3

US-09-819-964-5

Sequence 5, Application US/09819964

Patent No. 6369294

GENERAL INFORMATION:

APPLICANT: Piedrahita, Jorge A

TITLE OF INVENTION: Compositions and Methods for the Generation of Transgenic Animal Species

NUMBER OF SEQUENCES: 51

CORRESPONDENCE ADDRESS:

ADDRESSEE: ARNOLD, WHITE AND DURKEE

STREET: P.O. Box 4433

CITY: Houston

STATE: TX

COUNTRY: US

ZIP: 77210-4433

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/819, 964

FILING DATE: 28-Mar-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/949, 155

FILING DATE: <Unknown>

APPLICATION NUMBER: US 60/046, 094

FILING DATE: 09-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Hibler, David W.

REGISTRATION NUMBER: 41, 071

REFERENCE/DOCKET NUMBER: TANK:177

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (713) 789-2679

INFORMATION FOR SEQ. ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1126 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 51..1001

SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-819-964-5

Pred. No.:	2.79e-103	Length:	1126
Score:	1134.00	Matches:	223
Percent Similarity:	83.75%	Conservative:	45
Best Local Similarity:	69.69%	Mismatches:	44
Query Match:	71.55%	Indels:	8
	4	Gaps:	3

US-09-827-854-19 (1-317) x US-09-819-964-5 (1-1126)

```

Qy 1 MetysValleuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 51 ATGAGGGTTCTGTGGTGTCTTGTGTGTACCTCTCCGACGATGCGGACAGAGAGAC 110
Qy 21 GluGlnAlaValGluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 111 GAGCCGGGG-----CCGCGCGCGGAGGTGCAAGTGTGTGGAGAGAGAGCCCAAG 158
  
```



```

OY 38 TTPGLSerglyGlnArtrpGluLeuAlaLeuGlyArpPheTrpAspTyrLeuArtrp 57
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 159 TGGCAGGAGGAGCCGCTGGAGAGAGCCCTGGCCGCTTCTGGATTACCTGCCCTGG 218
OY 58 ValGlnThrLeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGlu 77
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 219 GTGCAGTCCCTGTCTACAGAGTGCAGAGAGAGCTCTCAGCACCAAGGACCCAGGAA 278
OY 78 LeuArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlu 97
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 279 CTGACGAGAGTGTATAGAGAGAGACATGAAGAGAGAGAGAGAGAGAGAGAGAGAG 338
OY 98 GlnGlnLeuThrProValAlaGlnGluThrArgAlaArgLeuSerLysGluLeuGlnAla 117
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 339 GCGCAGGAGGAGCCGCTGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 398
OY 118 AlaGlnAlaArgLeuGlnGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArg 137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 399 GCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 458
OY 138 GlyGlnValGlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSer 157
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 459 AGGAGAGTGCACAAATGTTGGGCCAGACACAGAGAGAGAGAGAGAGAGAGAGAGAG 518
OY 158 HisLeuArgLysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeu 177
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 519 CACCTTGGCAGAGCTGGCAGAGCGGCTGCTCCGCGACACCGAGAGAGAGAGAGAGAG 578
OY 178 AlaValTyrGlnAlaGlnAlaArgLysGlnGluArgGlnGlyLeuSerAlaIleArgGlu 197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 579 GCGGTATACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 638
OY 198 ArgLeuGlyProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAla 217
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 639 CCGCTGGGAGCCCTGTGTGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 698
OY 218 GlyGlnProLeuGlnGlnArgAlaGlnAlaIleArgGlyGlnArgLeuArgAlaArgMetGlu 237
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 699 GCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 758
OY 238 GlnMetGlySerArthrArgAspArgLeuAspGluValLysGlnGlnGlnValAlaGlnVal 257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 759 GAGATGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 818
OY 258 ArgAlaLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 277
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 819 CCGACCAAGTGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 878
OY 278 ArgLeuLysSerThrPheGluProLeuValGlnAspMetGlnArgGlnThrAlaGlyLeu 297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 879 CTCCTCAAGGCTGTTCAGAGCTCTGTGTGAGAGACATACGGGCGCAGTGGCGGCTGG 938
OY 298 ValGlnLysValGlnAlaValGly---ThrSerAlaAlaProValProSerAspAsn 316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 939 GTGGAGAGAGTGCAGTGGGCGGAGAGACATAGCTCTCCACCTCTGCGGCCAGTGAAT 998

```

RESULT 4
US-08-949-155-51

Sequence 51, Application US/08949155

Patent No. 6271436

GENERAL INFORMATION:

APPLICANT: Piedrahita, Jorge A

APPLICANT: Bazer, Fuller W

TITLE OF INVENTION: Compositions and Methods for the

NUMBER OF SEQUENCES: 51

CORRESPONDENCE ADDRESS:

ADDRESS: ARNOLD, WHITE AND DURKEE

STREET: P. O. Box 4433

CITY: Houston

STATE: TX

COUNTRY: US

ZIP: 77210-4433

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/949,155
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/027,338
; FILING DATE: 11-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,094
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hilder, David W.
; REGISTRATION NUMBER: 41,071
; REFERENCE/DOCKET NUMBER: TANK:177
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4267 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-949-155-51

```

Alignment Scores:
Pred. No.: 4,87e-89
Score: 998.00
Percent Similarity: 58.43%
Best Local Similarity: 48.76%
Query Match: 62.97%
DB: 4 Gaps: 5

```

US-09-827-854-19 (1-317) x US-08-949-155-51 (1-4267)
OY 1 MetLysValLeuThrAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    ||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2448 ATGGCTGTGTAAGTGGCTGGT-----GACAGATGCCGAGACAGAGAGC 2489
OY 21 GlnGlnAlaValGlnThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2490 GAGCGCGAG-----CCGCGCGGAGGAGTGCAGCTGTGTGGAGGAGGAGCCAG 2537
OY 38 TTPGLSerglyGlnArtrpGluLeuAlaLeuGlyArpPheTrpAspTyrLeuArtrp 57
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2538 TGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2597
OY 58 ValGlnThrLeuSerGluGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 77
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2598 GTGCAGTCCCTGTCTGACCAAGTGCAGAGAGAGAGCTGCTCAGCACCAAGGTCA 2657
OY 78 Leu----- 78
Db 2658 CT-GACGTAAGTGGCCACCCGAGCTCCCGCGCGCGCGCGCGCGCGCGCGCTGA 2716
OY 78 ----- 78
Db 2717 CCCTCTGGCGAAGCGGTGTCTGTGAGCCCTCAGCTCCACCCGCTTCCTTCG 2776
OY 78 ----- 78
Db 2777 TCCTTGTGCCACACTTGTGGGGGTCTGGGTCTGTCTTTCTTTTCTTCCTCTTT 2836
OY 78 ----- 78
Db 2837 TTGGGGGGAAGAACTTTTCTTTTCTTTTCTTTTCTTTTCACTTCATGTCTTCTTCC 2896
OY 78 ----- 78

```


D	2837	TTGGGGGAAAAAACATTTTCTTTCTTTCTTAATTGAATCATGTCCGTTCTTTC	2896
OY	78	-----	78
D	2897	ATCTTGAGCTCCTGCCTTCGCTGTCTCGGTCAAGCTTGCCGTCCTGTCTCTG	2956
OY	78	-----	78
D	2957	AATCTGACACGTCCTGGCCATCGCCAGCTCAGAGCCCTCTTCTCCCCCTCACGCC	3016
OY	79	-----ArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaIyr	92
D	3017	CCGCCCCCTCTCTGGCCACAGAGCGATNAGAAGAGAGCATAGAGGTGAAGSCCTAC	3076
OY	93	LysSerGluLeugluguInleuthrProValAlaGluThrArgAlaIrgLeuSer	112
D	3077	CGCAGAGAGCTGGAGGGCGAGCTGGGGCCCCGCTGACCAGAGAACCCAGCGCGCTGTCC	3136
OY	113	LysGluLeuglnAlaIaglnAlaArgLeuGlyAlaAspMetGluAspValCysLeIyrrg	132
D	3137	AAGAGCTGGAAGGGGGGCGAGGCCCGCGTGGGGCGGACATGAGAGAGCTGGCCAAAGC	3196
OY	133	LeuValIrgIntYrArgrglGylGluValAlGlnAlaMetLeuGlyGlnSerThrglugluleuArgr	152
D	3197	TGTGGTGCTACAGCGACAGAGAGGTGCACAACATGTTGGGCGACACACGAGAGAGCTGGG	3256
OY	153	ValArgLeuAlaSerHisLeuArqIyLysLeuArqGlnArqLeuLeuArqAspAlaAsp	172
D	3257	AGCGCGCTGGCTTCCCACCTGCCCAAGCTGCCAAGCGGCTCTCCGGACACCGAGGAC	3316
OY	173	LeuGlnIyrsArgLeuAlaValIyrglnAlaGlyAlaArgGluGlyAlaGlnAlaArgLeu	192
D	3317	CTGGAGAGGCGCTGGCGCTGTACTACAGGGCGGGCTGCGGAGGGGCGCGAGCGCAGCTG	3376
OY	193	SerAlaIleaRgtuarqLeuGlyProLeuValGlunGlnIyArgValArgAlaIarThr	212
D	3377	AGCGCCCTCGCGAGGCCCTCGGGCCCCCTGTGTGAGACAGGGCCGATTGCGCGCCGCCAC	3436
OY	213	ValGIserLeuAlaGlyGlnProLeuglngluArqAlaGlnAlaTrpGlyLuArgLeu	232
D	3437	CTGAGTACCAAGGGCGGGCCAGCGCGTGGCGAGCGGGCGGAAGCCTGGGGCCAGAACTG	3496
OY	233	ArgAlaIarmetGluGluMetGlySerArqThrArqAspArgLeuAspGluValLysGlu	252
D	3497	CGCGAGCGGTGAGGAGGAGGGCACCGAGCCCGGACCGCTGGAAGAGTGGCTGAG	3556
OY	253	GlnValIagIuValArqAlaLysleuglngluGlnAlaGlnGlnIlaArgLeuGlnAla	272
D	3557	CAGCTGGAGAGAGTGGCCACCAAGGAGGAGACAGGACGACCAATTTGGCTGTGAGGCC	3616
OY	273	GluAlaPheGlnAlaArgLeuLysSerTrpPheGluProLeuValGluAspMetGlnArq	292
D	3617	GAGGAGATTCCACGCGCTCCCAAAGCTGTGTGAGCGCTGTGTGGAAGACATACAGGGCC	3676
OY	293	GlnTrpAlaGlyLeuValGluLysValGlnAlaIalaValGly--ThrSerAlaAlaPro	311
D	3677	CAGTGGGCGCGGTGGTGGAGAGATGACGTGCGCGCTGACATTAAGCTCTCCACACTCT	3736
OY	312	ValProSerAspasn 316	
D	3737	GCGCCAGTGATAAT 3751	

US-08-726-306A-28
Sequence 28, Application US/08726306A
Patent No. 5958684
GENERAL INFORMATION:
APPLICANT: van Laeuwen, Frederik Willem
APPLICANT: Burscheld, Johannes Peter Henri
APPLICANT: Grosveld, Franklin G.
TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
NUMBER OF SEQUENCES: 189
CORRESPONDENCE ADDRESS:

	ADDRESS:	Panner & Wilcoff, Ltd.
	STREET:	1 Financial Center
	CITY:	Boston
	STATE:	MA
	COUNTRY:	US
	ZIP:	02111
	COMPUTER READABLE FORM:	
	MEDIUM TYPE:	Diskette, 3.50 inch, 1.44 Mb storage
	OPERATING SYSTEM:	PC-DOS/MS-DOS
	SOFTWARE:	WordPerfect 6.1
	CURRENT APPLICATION DATA:	
	APPLICATION NUMBER:	US/08/726,306A
	FILING DATE:	02-Oct-1996
	PRIOR APPLICATION DATA:	
	APPLICATION NUMBER:	GB 95/20080.4
	FILING DATE:	02-Oct-1995
	PRIOR APPLICATION DATA:	
	APPLICATION NUMBER:	US 60/009,832
	FILING DATE:	01-Jan-1996
	ATTORNEY/AGENT INFORMATION:	
	NAME:	Williams, Ph.D., Kathleen M.
	REGISTRATION NUMBER:	34,360
	REFERENCE/DOCKET NUMBER:	96,048-A (3255/00784)
	TELECOMMUNICATION INFORMATION:	
	TELEPHONE:	(617) 345-9100
	TELEFAX:	(617) 345-9111
	INFORMATION FOR SEQ ID NO:	28:
	SEQUENCE CHARACTERISTICS:	
	LENGTH:	660 base pairs
	TYPE:	nucleic acid
	STRANDEDNESS:	double
	TOPOLOGY:	linear
	MOLECULE TYPE:	cDNA
	US-08-726-306A-28	
	Alignment Scores:	
	Pred. NO.:	2.24e-88
	Score:	980.00
	Percent Similarity:	98.09%
	Best Local Similarity:	98.49%
	Query Match:	61.83%
	DB:	gaps: 2
		indels: 0
		mismatches: 0
		conservative: 1
		matches: 196
		length: 660
US-09-827-854-19 (1-317) x US-08-726-306A-28 (1-660)		
OY	1 MettysValLeutRPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal	20
Db		
Db	62 ATGAAGGTCTGTGGCGCTTCTGTCTGCATTTCTCTGCAGGATGCCAGGCCAAAGTG	122
OY	21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTPGInSer	40
Db		
Db	122 GAGCAAGCGGTGGAGACAGACC CGGAGCCGAGCTGGCGCCAGCAGACCGAGTGGCAAGC	181
OY	41 GlyGlnArgTPGIuLeuAlaLeuGlyAraPheTrpAspTryLeuArgTPyValGlnThr	60
Db		
Db	182 GCCAGCGCTGGGAATCGCAC TGGCTTTTGGGATTACTGCGCTGGGTGCAGACA	241
OY	61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80
Db		
Db	242 CTGTCTGACGAGGTGCGAGGAGGAGACTGCTCAGCTCCCAAGTCACCCAGAAGCATGAGGCG	301
OY	81 LeuMetAspGIuThrMetLysGluLeuLysAlaTyrllysSerGluLeuGlnGluGlnLeu	100
Db		
Db	302 CTGATGAGCAGACCATGTAAGAGATTGAAGCCCTTAATAATCGAATCGAGAACCAACTG	361
OY	101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIleGlnAla	120
Db		
Db	362 ACCCGGTAGCGGAGGAGACCGCGGCACGCTCTCCCAAGGAGCTGACAGCGCGCCAGGCC	421
OY	121 ArgLeuGlyAlaAspMetGluLysValCylserLysArgLeuValGlnlrrArgGlyGlnVal	140
Db		
Db	422 CGGCTGGGCGCGGACATGAGACAGTGTGGCGCGCTGTGGTGCATGCCCGGCGAGGTG	481

QY 141 GlnAlaMetIeuGlnGlnSerThgGluGluLeuArgAlaArgLeuAlaSerHisIleuArg 160
Db 482 CAGGCCATGCTCGGCCAGACGACCGAGAGAGCTGGGGGTGGCCCTCGCTCCACCTTCGC 541
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnIlyAsrArgLeuAlaValTyr 180
Db 542 AAGCTGCGCTAAGCGGCTCTCTCGGGATTCGCGATACCTGCGAAGCGCCCTGGCAGTGTAC 601
QY 181 GlnAlaGlnAlaArgGlnGlnAlaGlnAlaArgGlnLysLeuSerAlaIleArgGlnArgLeu 199
Db 602 CAGCGCGGGGCGCCGAGAGGCGCCGAGACGCGCGCTCAAGCGCATCCCGAGCGGCTG 658

QY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
|||||
Db 2 GCACGGCTGCTCCAAAGAGCTGCAGCGCGCAGCGCCGCGCTGCGCGCGCATGAGAGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
|||||
Db 62 GGTGTGGC-CGCTGTGTGAGTACCGCGCGAGTGTACAGCCATGCTCGCGCAGACAC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgGlnArgLeuLeuArg 168
|||||
Db 121 GAGGAGCTCGGGTGGCTCGCTCGCTCCACCTGCGCAAGCTGTGAAGCGGCTCTCCCG 180
QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
|||||
Db 181 GATGCCGATGACTGCGAAGATGCTGTGCGAGTGTACAGCGCGCGCGCGCGAGGCGCC 240
QY 189 GluArgGlyLeu 192
|||||
Db 241 GAGCGCGGCTC 252
RESULT 9
US-09-287-141-24
; Sequence 24, Application US/09287141
; Patent No. 6197498
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037-9103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/287,141
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 18-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/406,199
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-2002D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-587-5360
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-287-141-24
Alignment Scores:
Pred. No.: 12e-28 Length: 252
Score: 370.00 Matches: 81

Percent Similarity: 97.628 Conservative: 1
Best Local Similarity: 96.438 Mismatches: 2
Query Match: 23.348 Indels: 1
DB: 4 Gaps: 0
US-09-827-854-19 (1-317) x US-09-287-141-24 (1-252)
QY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
Db 2 GCACGGCTGCTCCAAAGAGCTGCAGCGCGCAGGCTGCGCGCATGAGAGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
|||||
Db 62 GGTGTGGC-CGCTGTGTGAGTACCGCGCGAGTGTACAGCCATGCTCGCGCAGACAC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgGlnArgLeuLeuArg 168
|||||
Db 121 GAGGAGCTCGGGTGGCTCGCTCGCTCCACCTGCGCAAGCTGTGAAGCGGCTCTCCCG 180
QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
|||||
Db 181 GATGCCGATGACTGCGAAGATGCTGTGCGAGTGTACAGCGCGCGCGCGAGGCGCC 240
QY 189 GluArgGlyLeu 192
|||||
Db 241 GAGCGCGGCTC 252
RESULT 10
US-09-431-613-24
; Sequence 24, Application US/09431613
; Patent No. 6221601
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037-9103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 18-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/406,199
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-2002G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-587-5360
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

MOLECULE TYPE: CDNA
US-09-431-613-24

Alignment Scores:

Pred. No.:	1.2e-28	Length:	252
Score:	370.00	Matches:	81
Percent Similarity:	97.62%	Conservative:	1
Best Local Similarity:	96.43%	Mismatches:	2
Query Match:	23.34%	Indels:	1
DB:	4	Gaps:	0

US-09-827-854-19 (1-317) x US-09-431-613-24 (1-252)

QY 109 AlaArgLeuSerLysGluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGCGCGCGCGCGCGCTGGCGGACATGAGAGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTCCGC-CGCTGTGTGAGTACCGCGCGCGCGCATGCTCGCGCAGAGAC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgGlnArgLeuLeuArg 168
DB 121 GAGGAGCTGGGGGTGGCTGCTCCACCTGCGCAAGCTGTAAGCGGCTCTCCGC 180

QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGTCCCTGGCAGTGTACAGCGCGCGCGCGCGAGGCGCC 240

QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTC 252

RESULT 11
US-09-504-245-24
Sequence 24, Application US/09504245
Patent No. 6221605
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe LLP
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/504,245
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002J
TELECOMMUNICATION INFORMATION:
TELEPHONE: 858-450-8400
TELEFAX: 858-587-5360
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-504-245-24

Alignment Scores:

Pred. No.:	1.2e-28	Length:	252
Score:	370.00	Matches:	81
Percent Similarity:	97.62%	Conservative:	1
Best Local Similarity:	96.43%	Mismatches:	2
Query Match:	23.34%	Indels:	1
DB:	4	Gaps:	0

US-09-827-854-19 (1-317) x US-09-504-245-24 (1-252)

QY 109 AlaArgLeuSerLysGluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGCGCGCGCGCGCTGGCGGACATGAGAGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTCCGC-CGCTGTGTGAGTACCGCGCGCGCGCGCATGCTCGCGCAGAGAC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgGlnArgLeuLeuArg 168
DB 121 GAGGAGCTGGGGGTGGCTGCTCCACCTGCGCAAGCTGTAAGCGGCTCTCCGC 180

QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGTCCCTGGCAGTGTACAGCGCGCGCGCGCGAGGCGCC 240

QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTC 252

RESULT 12
US-09-287-682-24
Sequence 24, Application US/09287682
Patent No. 6235478
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002E
TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-682-24

Alignment Scores:
Pred. No.: 1.2e-28 Length: 252
Score: 370.00 Matches: 81
Percent Similarity: 97.628 Conservative: 1
Best Local Similarity: 96.43% Mismatches: 2
Query Match: 23.34% Indels: 1
DB: Gaps: 0

US-09-827-854-19 (1-317) x US-09-287-682-24 (1-252)

QY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
|||||
DB 2 GCACGGCTGTCCAAAGAGCTGCAGCGCGCGCAGCCCGCTGGCGCGGACATGAGAGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
|||||
DB 62 GTGTGGCG-CGGCTGTGTGAGTACCGCGCGGAGGTGCAGGCCCTGTGGCCGAGACACC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgGlnArgLeuLeuArg 168
|||||
DB 121 GAGGAGCTCGGGGTGGCTCGCTCCGCTCCACCTGCGCAGCTGCTGAAGCGCTCTCTCCGC 180

QY 169 AspaAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
|||||
DB 181 GATGCCGATGACTCGCAGAAATCCTCGCAGTGTACACGAGCGCGCGCGCGAGGCGCC 240

QY 189 GluArgGlyLeu 192
|||||
DB 241 GAGCGCGGCTC 252

RESULT 13
US-09-287-679-24
; Sequence 24, Application US/09287679
; Patent No. 6258538
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
APPLICANT: Braun, Andreas
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,679
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-679-24

Alignment Scores:
Pred. No.: 1.2e-28 Length: 252
Score: 370.00 Matches: 81
Percent Similarity: 97.62% Conservative: 1
Best Local Similarity: 96.43% Mismatches: 2
Query Match: 23.34% Indels: 1
DB: Gaps: 0

US-09-827-854-19 (1-317) x US-09-287-679-24 (1-252)

QY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
|||||
DB 2 GCACGGCTGTCCAAAGAGCTGCAGCGCGCGCAGCCCGCTGGCGCGGACATGAGAGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
|||||
DB 62 GTGTGGCG-CGGCTGTGTGAGTACCGCGCGGAGGTGCAGGCCATGCTCGCGCAGACACC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgGlnArgLeuLeuArg 168
|||||
DB 121 GAGGAGCTCGGGGTGGCTCGCTCCGCTCCACCTGCGCAGCTGCTGAAGCGCTCTCTCCGC 180

QY 169 AspaAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
|||||
DB 181 GATGCCGATGACTCGCAGAAATCCTCGCAGTGTACACGAGCGCGCGCGAGGCGCC 240

QY 189 GluArgGlyLeu 192
|||||
DB 241 GAGCGCGGCTC 252

RESULT 14
US-09-397-766-24
; Sequence 24, Application US/09397766
; Patent No. 6268144
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
APPLICANT: Braun, Andreas
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/397,766
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-20021
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-397-766-24

Alignment Scores:
Pred. No.: 1,2e-28 Length: 252
Score: 370.00 Matches: 81
Percent Similarity: 97.62% Conservative: 1
Best Local Similarity: 96.43% Mismatches: 2
Query Match: 23.34% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-19 (1-317) x US-09-397-766-24 (1-252)

OY 109 AAlaArgLeuSerLySGluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGlnuSp 128
DB 2 GCACGGCTGTCCAAAGAGCTGCAGGCGGCGGCCCGCTGGCGGACATGGAGGAC 61
OY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTCCGC-CGCTGTGTGAGTACCGGGCGAGGTGCAGGCCATCTCGGCCAGACACC 120
OY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgGlnArgLeuArg 168
DB 121 GAGGAGCTGCGGGTCCGCTCCCTCCACCTGCGCAAGCTGCTAAAGCGGCTCTCCGC 180
OY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGlnGlyAla 188
DB 181 GATGCCGATGACCTGCAGAACTCCTGGCAGTGTACCAAGCGGGCCCGCGAGGGCGCC 240
OY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTC 252

RESULT 15

US-09-287-681-24
Sequence 24, Application US/09287681
Patent No. 6277573

GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,681
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-287-681-24

Alignment Scores:
Pred. No.: 1,2e-28 Length: 252
Score: 370.00 Matches: 81
Percent Similarity: 97.62% Conservative: 1
Best Local Similarity: 96.43% Mismatches: 2
Query Match: 23.34% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-19 (1-317) x US-09-287-681-24 (1-252)

OY 109 AAlaArgLeuSerLySGluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGlnuSp 128
DB 2 GCACGGCTGTCCAAAGAGCTGCAGGCGGCGGCCCGCTGGCGGACATGGAGGAC 61
OY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTCCGC-CGCTGTGTGAGTACCGGGCGAGGTGCAGGCCATCTCGGCCAGACACC 120
OY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgGlnArgLeuArg 168
DB 121 GAGGAGCTGCGGGTCCGCTCCCTCCACCTGCGCAAGCTGCTAAAGCGGCTCTCCGC 180
OY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGlnGlyAla 188
DB 181 GATGCCGATGACCTGCAGAACTCCTGGCAGTGTACCAAGCGGGCCCGCGAGGGCGCC 240
OY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTC 252

Search completed: March 14, 2003, 20:18:56
Job time : 39.7343 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Comphen Ltd.

OM protein - nucleic acid search, using frame_plus.p2n model

Run on: March 14, 2003, 17:32:58 ; Search time 65.8753 Seconds
(without alignments)
3377.460 Million cell updates/sec

Title: US-09-827-854-19

Perfect score: 1585
Sequence: 1 MKVLMALLVTFAGCCQAKV.....VEKVOAAVGTSAAPVSDNH 317

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 501302 seqs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus.p2n.model -DEV=xlp
-CGN2_1/USPTO_SPOOL/US09827854/runat_11032003.101612.27578/app.query.fasta.1.3576
-DB=Published_Applications_NA -OFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=DLSUM62
-THR_MIN=0 -ALIGN=45 -DOCALIGN=200 -THR_SCORE=pcr -THR_MAX=100 -MINLEN=0
-MAXLEN=200000000 -USER=US09827854.lecgn1.1.298.runat_11032003.101612.27578
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MAP -LARGEDEPT -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -MARK_TIMEOUT=30 -THREADS=1 -XGAPOP=1 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:*

1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubna/PCF_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubna/PCFUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1585	100.0	1156	10	US-09-827-854-12
2	1581	99.7	1156	9	US-09-870-759-129
3	1581	99.7	1156	9	US-09-802-640-17
4	1581	99.7	1156	10	US-09-827-854-8

	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
	1581	1573	1573	1572	1572	1569	1566	1365.5	867.5	625	625	459	458.5	439.5	435	428.5	422.5	405.5	401.5	397	387.5	380.5	379.5	379.5	379.5	375.5	370	370	369.5	369.5	351.5	350.5	350.5	350.5	348.5	345.5	344.5	338.5	338	337.5	335.5
	99.7	99.2	99.2	99.2	99.2	99.0	98.8	1365.5	54.7	39.4	39.4	29.0	28.9	27.7	27.4	27.0	26.7	25.6	25.3	25.0	24.4	24.0	23.9	23.9	23.9	23.7	23.3	23.3	23.3	23.3	22.2	22.1	22.1	22.1	22.0	21.8	21.7	21.4	21.3	21.3	21.2
	1291	1156	1156	1157	1157	1156	1156	41907	786	478	478	356	449	442	414	416	423	409	416	425	425	390	391	392	401	410	410	410	404	405	377	353	386	388	396	377	376	332	376	373	
	US-10-044-090-454	US-09-827-854-9	US-09-827-854-11	US-09-954-456-760	US-09-880-107-2244	US-09-827-854-7	US-09-827-854-10	US-09-967-013-5	US-09-925-302-133	US-09-964-824-374	US-09-880-107-2491	US-09-960-352-5420	US-09-960-352-4726	US-09-960-352-1395	US-09-960-352-8237	US-09-960-352-8720	US-09-960-352-14047	US-09-960-352-5148	US-09-960-352-9797	US-09-179-5368-130	US-09-960-352-3497	US-09-960-352-1311	US-09-960-352-1278	US-09-960-352-10599	US-09-960-352-1187	US-09-960-352-8042	US-09-796-416-24	US-09-879-341-24	US-09-960-352-2325	US-09-960-352-14063	US-09-960-352-176	US-09-960-352-4914	US-09-960-352-10040	US-09-960-352-10361	US-09-960-352-11966	US-09-960-352-1457	US-09-960-352-10540	US-09-960-352-8934	US-09-960-352-6552	US-09-960-352-5408	US-09-960-352-3681
	Sequence 454, App	Sequence 9, Appl	Sequence 11, Appl	Sequence 760, App	Sequence 2244, App	Sequence 7, Appl	Sequence 10, Appl	Sequence 5, Appl	Sequence 133, App	Sequence 2491, App	Sequence 5420, App	Sequence 4726, App	Sequence 9959, App	Sequence 4337, App	Sequence 8720, App	Sequence 14047, A	Sequence 5148, App	Sequence 9797, App	Sequence 130, App	Sequence 3497, App	Sequence 1311, App	Sequence 1278, App	Sequence 10599, A	Sequence 1187, App	Sequence 8042, App	Sequence 24, Appl	Sequence 24, Appl	Sequence 2225, App	Sequence 14063, A	Sequence 176, App	Sequence 4914, App	Sequence 10040, A	Sequence 10361, A	Sequence 11966, A	Sequence 1457, App	Sequence 10540, A	Sequence 8934, App	Sequence 6552, App	Sequence 5408, App	Sequence 3681, App	

ALIGNMENTS

RESULT 1
US-09-827-854-12
; Sequence 12, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassilis
; APPLICANT: Kyriacos, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; TITLE OF INVENTION: Cholesterol levels without inducing
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-12

Alignment Scores:
Pred. No.: 1.67e-133 Length: 1156
Score: 1585.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

OY		1	MetylsValLeuTrpAlaIleuLeuValThrPheLeuAlaGlyCysGlnAlaIleVal	20
Db		61	ATGAAGGTCTGTGGCTGCCCTTCTGTGCATTCTCGGCAGATGCCAGGCCAAGTC	120
OY		21	GIuGlInAlaValGIuThrGIuProGIuProGIuLeuArgInGIuThrGIuTrpGIuSer	40
Db		121	GAGCAACCGGGTGAGACAGAAGCCGAGACCAGCTGGCCAGACAGAACGAGTGGACAGC	180
OY		41	GLyLIuArgrTrpGIuLeuAlaIleuGlyArgPheTrpAspIyrLeuArgrTrpValGIuThr	60
Db		181	GGCCAGCGCTGGGAACAGGCACCTGGGTCTGTTGGATTACTGGCGCTGGGTGCAGACA	240
OY		61	LeuSerGIuGIuValGIuGIuGIuLeuLeuSerSerGIuValThrGIuGIuLeuArgAla	80
Db		241	CTGTCTAGACAGGTGCAGAGAGAGACTCTCAGCTCCAGGTCACCCAGAACTCAGAGCGC	300
OY		81	LeuMetAspGIuThrMetyLysGIuLeuLysAlaTrpLysSerGIuLeuGIuGIuLeu	100
Db		301	CTGATGGCGAGACCATGAAGAGATTGAAGCCCTCAATAATCGGAATCGTAGAGAACTC	360
OY		101	ThrpProValAlaGIuGIuThrArgAlaArgLeuSerLysGIuLeuGIuAlaIleuAla	120
Db		361	ACCCCCGGTGGGAGAGAGACCGGGGACGGGTCTGCCAAGAGCTGCAGGGGGCGAGCC	420
OY		121	ArgLeuGIuAlaAspMetGIuAspValCysGIuArgLeuValGIuTrpArgGIuVal	140
Db		421	CGCGTGGGGCGGACATGAGAGAGACTGTGGGGCGCCCTGGTGCAGTACCGCGGAGGTC	480
OY		141	GIuAlaMetLeuGIuGIuSerThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg	160
Db		481	CAGGCCATGTCTCGGCCAGACACCCGAGAGACTGGGGTGGCCCTCGCTCCCACCTGGCC	540
OY		161	LysLeuArgGIuArgLeuLeuArgAspAlaAspAspLeuGIuLysArgLeuAlaValI Tyr	180
Db		541	AAGCTGGCTCAGCGGCTCTCTCCGCGATGCCGATGACTCTGCAGAACGCCCTGGCGATGAC	600
OY		181	GIuAlaGIuAlaArgGIuGIuAlaGIuArgGIuLysSerAlaIleArgLeuArgLeuGIu	200
Db		601	CAGCGCGGGGCCCGAGAGGGCGCCGAGCGCGGCTCAGCGCCATCCGCGAGCGCGCTGGGG	660
OY		201	ProLeuValGIuGIuGIuArgValArgAlaAlaI ThrValIeLysSerLeuAlaGIuGIuPro	220
Db		661	CCCCGTGTGGAAACGGGCGCGCTGGCGGCGCACATCTGGGGCTCTCGCGCGGCCACGCCG	720
OY		221	LeuGIuGIuArgAlaGIuAlaI TrpGIuLysArgLeuArgAlaTrpMetGIuGIuMetGIu	240
Db		721	CTACAGAGAGCGGGGCCACGAGCCTGGGGGAGACGGCTGGCGCGCGGATGAGAGAGATGGG	780
OY		241	SerArgThrArgAspArgLeuAspGIuValLysGIuGIuAlaIleuAlaGIuValArgAlaLys	260
Db		781	AGCGGAGCCCGCGACCCGCTGGAGACAGAGTAAGAGAGCAGGTGGCGGGCGCGCGCAAG	840
OY		261	LeuGIuGIuGIuAlaGIuGIuGIuIleArgLeuGIuAlaGIuAlaPheGIuAlaArgLeuLys	280
Db		841	CTTGAGAGAGAGCGCCACAGCATACGCTCTCAGGCCGAGCCCTTCCAGCGCCCGCTTCAAG	900
OY		281	SerTrpPheGIuProLeuValGIuAspMetGIuArgGIuTrpAlaGIuLeuValGIuLys	300
Db		901	AGCTGGTTCAGGCCCTGGTGGAGAGACATCAGCCCAATGGGGCCGGGCTGGTGGAGAAAG	960
OY		301	ValGIuAlaAlaValGIuTrpSerAlaAlaProValProSerAspAsnHis	317
Db		961	GTGCAGGCTGCGTGGGACACAGCGCGCCCTGTGGCCAGCAATATC	1011
RESULT 2				
US-09-870-759-129				
Sequence 129, Application US/09870759				

```

? Patent NO.: US20020177551A1
? GENERAL INFORMATION:
? APPLICANT: TERMAN, David S
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
? FILE REFERENCE: 870759
? CURRENT APPLICATION NUMBER: US/09/870,759
? CURRENT FILING DATE: 2002-01-14
? PRIOR APPLICATION NUMBER: US 60/208,128
? PRIOR FILING DATE: 2000-05-30
? NUMBER OF SEQ ID NOS: 166
? SOFTWARE: Patentin version 3.1
? SEQ ID NO: 129
? LENGTH: 1156
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (61)..(1014)
? OTHER INFORMATION:
? US-09-870-759-129

```

Pred. No.:	3.8e-133	Length:	1156
Score:	1581.00	Matches:	316
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.68%	Mismatches:	0
Query Match:	99.75%	Indels:	0
DB:	9	Gaps:	0
US-09-827-854-19 (1-317) x US-09-870-759-129 (1-1156)			
Oy	1 MetlysvallentrrpalalaleuleuvalthrPheleuaactlysglnalalyval	20	
Dd	ATGAAGGTTCTGTGGCTGGCTTGCTCGTACATTCTCTGCAGAGATCCAGGCCAAGGTG	120	
Oy	21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer	40	
Dd	GAGCAAGCGGGGAGACAGAGACCAGCAGCTGGCCAGCAGACGAGCCGAGTGGCAGAGC	180	
Oy	41 GlyGlnArgrTPrgluleuAlaLeuGlArGrPhetrPaSPyrleuArgrTPaValGlnThr	60	
Dd	GGCCAGCGCTGGAACTGGCACCTGGTCGCTTTGGATTACTCGCGCTGGTGGACAGA	240	
Oy	61 LeuSerGlnGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80	
Dd	CTGTCTAGCAGCGTGCAGAGAGACTCTCAGCTCCACAGCTCACCCAGCAACTGAGAGGGC	300	
Oy	81 LeuMetaspGluThrMetLysGluLeuLysAlaThyTylSerGluLeuGlnGlnGlnLeu	100	
Dd	CTGTATGGACGACCATGTAAGAGGATTGAAGGCTTACAATAATCGGAACCTGAGAACACTG	360	
Oy	101 ThrProValAlaGlnGluThrArgAlaArGrLeuSerLysGluLeuGlnAlaAlaGlnAla	120	
Dd	ACCCCGGTGGCGAGAGACCGCGGGCACGGCTGTCCAAGAGACTGCAGCGGGCGCAGGCC	420	
Oy	121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnThrArgGlyGluVal	140	
Dd	CGGCTGGGGCGGACATGGAAGAGTGTGGCGGCGCTGTGCAGTACCGCGGCGAGGTG	480	
Oy	141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuLeuArgValArgLeuAlaSerHisLeuArg	160	
Dd	CAGCCATGTCTCGGCCAGACACCGAGAGACTGGGGTGGCTCGCTGCCACCTGGCGC	540	
Oy	161 LysLeuArgrGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValThr	180	
Dd	AAGCTGGGTAAAGCGGCTCTCCCGGATGCCGATGACTGCGAAGAGCGCTGGCAGTGTAC	600	
Oy	181 GlnAlaGlyAlaArgrGlnGlyAlaGluArgGlyLeuSerLysIleArGrGlnArgLeuGly	200	
Dd	CAGGCGCGGGCGCGCGAGGGCGCGAGCGGCTCAGCGCCATCCGCGAGCGGCTGGGG	660	
Oy	201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro	220	

Db 661 CCCCTGTGTGAACAGGCGCGGTGGCGCCCACTGTGGGCTCCCTGGCCGCGCAGCCG 720
Qy 221 LeuGlnGluArgAlaGlnAlaTyrPgluArgLeuArgAlaArgMetGluGluMetGly 240
Db 721 CTACAGGAGCGGCGCCAGCCCTGTGGGCGAGCGGTGGCGCGCGGATGAGAGATTGGGC 780
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGlnValArgAlaLys 260
Db 781 AGCGGAGCCCGGACCGGCTGTGGAGAGTGAAGAGAGAGTGGCGAGAGTGGCGCCAG 840
Qy 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 841 CTGGAGGAGCAGCGCCAGACAGATACGCTTGAGCGCGGCTTCAAGCGCGCTCAAG 900
Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 901 AGCTGTGTGAGCCCTGTGTGGAAGACATGACAGCCAGTGGCGGCGGTGTGGAGAG 960
Qy 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCGGTGGGACAGCGCGCCCTGTGGCCAGCAGCAATCAC 1011

RESULT 3
US-09-802-640-17
: Sequence 17, Application US/09802640
: Publication No. US20030036057A1
: GENERAL INFORMATION:
: APPLICANT: Braun, Andreas
: APPLICANT: Bonnal Aruna
: APPLICANT: Kleyn Patrick
: TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND THEIR USE
: FILE REFERENCE: 24736-2048
: CURRENT APPLICATION NUMBER: US/09/802,640
: CURRENT FILING DATE: 2001-03-09
: NUMBER OF SEQ ID NOS: 122
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 17
: LENGTH: 1156
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (61)...(1014)
: OTHER INFORMATION: Nucleotide sequence encoding apolipoprotein E
: OTHER INFORMATION: (APOE)
US-09-802-640-17

Alignment Scores:
Pred. No.: 3.8e-133 Length: 1156
Score: 1581.00 Matches: 316
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.68% Mismatches: 0
Query Match: 99.75% Indels: 0
DB: Gaps: 0

US-09-827-854-19 (1-317) x US-09-802-640-17 (1-1156)

Qy 1 MetLysValLeuTrpPheAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGGTTCTGTGGCGCTGTCTGTGTCAATTCCTGCAGATGCCAGGCCAAGGTG 120
Qy 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAGAGGGGTGAGACAGACCGGAGCCGAGCTGCGCCAGCAGACGAGTGGCAGAGC 180
Qy 41 GlyLysArgTrpGluLeuAlaLeuGlyLysArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 181 GGCCAGCGCTGGGAACATGCGATGGGTGCTTTGGGATTAACCTGCGGTGGGTGAGACA 240
Qy 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGCTGTGAGCAGGTGACAGGAGAGCTGCTCAGTCCAGAGTCCACCGAAGAACTGAGGCG 300

Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
Db 301 CTGATGAGAGACCATTAAGAGATTGAAGCCTTCAATGAGATGAGATGAGAGAACACTG 360
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 361 ACCCGGTGGGAGAGAACAGCGGCGACAGCTGTCTCAAGAGATGACAGCGGCGAGCC 420
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 421 CGGCTGGGCGGAGCATGAGAGAGCTGTGGCGCGCTGTGTGACATGCGGCGAGGTG 480
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCGATGTCTGGCCAGAGACACGAGAGAGCTGGGGTGGCTGTGCTCCACCTGGCG 540
Qy 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTCGCTAAGCGGCTCTCCGGATGCCATGACCTGCAAGAGCGCTGGCAGTGTAC 600
Qy 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 601 CAGCGCGGCGCCCGAGGCGCGAGCGCGCTCAAGCGCCATCCGAGCGCGCTGGGG 660
Qy 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGTGTGAACAGGCGCGGTGGCGCCCACTGTGGGCTCCCTGTGGCGCGCAGCG 720
Qy 221 LeuGlnGluArgAlaGlnAlaTyrPgluArgLeuArgAlaArgMetGluGluMetGly 240
Db 721 CTACAGGAGCGGCGCCAGCCCTGTGGGCGAGCGGTGGCGCGGATGAGAGATGGCG 780
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGlnValArgAlaLys 260
Db 781 AGCGGAGCCCGGACCGGCTGTGGAGAGTGAAGAGAGAGTGGCGGAGTGGCGCCAG 840
Qy 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 841 CTGAGAGAGAGCGCCAGACAGATACGCTTGAGCGCGGCTTCCAGGCGCGCTCAAG 900
Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 901 AGCTGTGTGAGCCCTGTGTGGAAGACATGACAGCGCCAGTGGCGGCGGTGTGGAGAG 960
Qy 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCGGTGGGACAGCGCGCCCTGTGGCCAGCAGCAATCAC 1011

RESULT 4
US-09-827-854-8
: Sequence 8, Application US/09827854
: Patent No. US20020123093A1
: GENERAL INFORMATION:
: APPLICANT: Zannis, Vassilis
: APPLICANT: Kyriacs, Vassilis
: TITLE OF INVENTION: Compounds and methods for lowering
: TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia
: FILE REFERENCE: 07180/004003
: CURRENT APPLICATION NUMBER: US/09/827,854
: CURRENT FILING DATE: 2001-04-05
: PRIOR APPLICATION NUMBER: US 09/679,088
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 09/544,386
: PRIOR FILING DATE: 2000-04-06
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 8
: LENGTH: 1156
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-827-854-8


```
Dh 756 CAGGCCGGGGCCCCGAGGGGCCGACGCGGCTCAGCGCCATCCGCGAGCCGCTGGG 785
Qy 201 ProleuValAGluGluGlyArgValArgAlaIaThrValGlySerLeuAlaGluGlnPro 220
Db 786 CCCCTGGTGGAAAGAGGCGGCGGGGCGCCACTGTGGCTCCCTGGCCGACGCG 845
Qy 221 LeuGluGluArgAlaGluAlaIaThrPslGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
Db 846 CACAGAGAGCGGGGCCGAGGCGCTGGGGCCAGCGGCTGGCGGCGGATGGAGAGATGGGC 905
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 906 ACCCGAGCCCGGACCGCGCTGGTGGACGAGTGAAGAGAGAGTGGCGAGGTGGCGCCCAAG 965
Qy 261 LeuGluGluGlnAlaGluGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db 966 CTGGAGAGAGAGCCGACGAGATGACGCTCGAGCGCGAGGCTTCACAGGCCGCGCTCAAG 1025
Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 1026 ACCTGGTTCGAGCCCTGGTGGACACATGCAGCCGCAATGGGCGGCGGTGGTGGAGAG 1085
Qy 301 ValGlnAlaIaValAGlyThrSerAlaIaProValProSerAspAsnHis 317
Db 1086 GTGCAAGGCTGGCGTGGGACACGCGCCGCTGTGCCACGACATCAC 1136

RESULT 6
US-09-827-854-9
; Sequence 9, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassilis
; APPLICANT: Kyriacos, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827, 854
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679, 088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544, 386
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-9

Alignment Scores:
Pred. No.: 1.98e-132 Length: 1156
Score: 1573.00 Matches: 315
Percent Similarity: 99.68% Conserves: 1
Best Local Similarity: 99.37% Mismatches: 1
Query Match: 99.24% Indels: 0
DB: 10 Gaps: 0

US-09-827-854-19 (1-317) x US-09-827-854-9 (1-1156)
Qy 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGTTCTGTGGCGTGGCTGCTGTCACATTCCTGGCGAGATGCGACGCCCAAGTG 120
Qy 21 GluGlnAlaValAGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTGCCGACAGACAGCGAGTGGCAGAGC 180
Qy 41 GlyLysArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLysArgTrpValGlnThr 60
Db 181 GCCCAGCGCTGGAGAACTGGCACTGGGTGGATTTGGGATTTACTCGCTGGGTGGACACA 240
Qy 61 LeuSerGluGlnValAGluGluGlnIleuLeuSerSerGlnValThrGlnGluLeuArgAla 80
```

```
Db 241 CTGTGAGACAGGTGACAGAGAGAGGCTGCTCACCTCCAGAGTCAACCAAGAACTGAGGGCG 300
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu 100
Db 301 CTGATGGACGAGACCATGATGAAGAGTTGAAGGCTCAAAATCGAACTGGAGGACAACAG 360
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaGlnAla 120
Db 361 ACCCGGTGGCGGAGAGAGAGCGGCGACGCTGTCCAAAGAGACTCAAGCGCGAGCGC 420
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGluTrpArgLysGluVal 140
Db 421 CGCGTGGGCGCGACATGAGAGAGAGCTGTGGCGCCCTGTGTCTGCTGCTGCTGCTGCTG 480
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTCGGCGACAGACCGAGAGAGTGGCGGTGCGCTGCTGCCACCTGGCGC 540
Qy 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTGCTGTAAGCGCTCTCCGCGATGCCATGACCTGACAGAACTGCTGGCAGTGTAC 600
Qy 181 GlnAlaGlyAlaArgGluGlyValAGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 601 CAGGCCGGGGCGCGGAGAGCGCGCAGCGCGGCTCAGAGCCATCGCGAGCGCTGGGG 660
Qy 201 ProLeuValAGluGluGlyArgValArgAlaIaThrValGlySerLeuAlaGluGlnPro 220
Db 661 CCCCTGGTGGAAAGAGGCGCGCTGGGCGGCGCTGTGGCTCCCTGGCCGCGGACGCG 720
Qy 221 LeuGluGluArgAlaGluAlaIaThrPslGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
Db 721 CTACAGAGAGCGGCGGAGGCGCTGGGCGGAGCGGCTGGCGGCGGATGGAGAGATGGGC 780
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 781 AGCGGAGACCGCGGACCGCTGGGCGGAGGAGGAGAGAGAGTGGCGGAGGTGGCGCCCAAG 840
Qy 261 LeuGluGluGlnAlaGluGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db 841 CTGGAGAGAGAGCGGCGGACGAGATGCGCTGCGAGCGCGGCTTCACAGGCCGCTCAAG 900
Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 901 ACCTGGTTCGAGCCCTGGTGGAGACATGACGCGCCAGTGGGCGGCGGTGGTGGAGAG 960
Qy 301 ValGlnAlaIaValAGlyThrSerAlaIaProValProSerAspAsnHis 317
Db 961 GTGCAAGGCTGGCGTGGGACACGCGCCGCTGTGCCACGACATCAC 1011

RESULT 7
US-09-827-854-11
; Sequence 11, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassilis
; APPLICANT: Kyriacos, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827, 854
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679, 088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544, 386
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
```


US-09-827-854-11

Alignment Scores:

Pred. No.:	1,98e-132	Length:	1156
Score:	1573.00	Matches:	315
Percent Similarity:	99.68%	Conservative:	1
Best Local Similarity:	99.37%	Mismatches:	1
Query Match:	99.24%	Indels:	0
DB:	10	Gaps:	0

US-09-827-854-19 (1-317) x US-09-827-854-11 (1-1156)

```
OY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB ATCAAGGTTCTGTGGCTGCTGTGTCACATTCCTGACAGATGCCAGGCAAGGTG 120
OY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrPgiInser 40
DB 121 GAGCAAGCGGTGAGACAGCGCGAGCCGAGCTCGCCAGCAGCGAGCGAGTGGCAGAGC 180
OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 181 GGGCAGCGCTGGGAAGTGGACACTGGCTGCTTTGGGATTACCTGGCGTGGTGACAGACA 240
OY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 241 CTGTCTGACAGACAGTGCAGAGAGCTGCTCAGCTCCAGGTCAACCAGAACTGAGGGCG 300
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu 100
DB 301 CTGATGACAGACCATGAAAGAGTTGAAGCCTACAAATCGGAATGGAGGGAACACTG 360
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 361 ACCCGCGTGGCGAGAGACGCGCGGACGGCTGTCCAGAGCTGACAGGGCGCGCAGGCC 420
OY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 421 CGGCTGGCGCGACATGAGAGACGTGTGCGCGCTGTCAGTACGCGCGCGAGAGTG 480
OY 141 GlnAlaMetLeuGlyGlnSerThrGluGluArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGCCATGCTGGCCAGACACCGAGAGACTGCGGGTGGCTCCCTCCACCTGCGCG 540
OY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 541 AAGCTGTAAAGCGCTCTCCCGCGATGCCGATGACCTGCAGAAAGCGCTGGCAGAGTAC 600
OY 181 GlnAlaGlyAlaArgGluGluGluArgGlyLeuSerAlaLeuArgGluGluGluGly 200
DB 601 CAGCGCGGGGCGCGAGGCGCGCGCGCTCAGCGCCATCCGCGAGCGCGCTGGGG 660
OY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 661 CCCCTGTGTGAACAGGCGCGCGCGCGCGCTGTGGCTCCCTGGCGCGCGCGCGCG 720
OY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluGly 240
DB 721 CTACAGAGACGCGCGCGCGCGCTGGCGCGAGCGCTGCGCGCGGATGAGAGATGGGG 780
OY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 781 ACCCGGACCGCGCGCGCTGCGAGGAGTGAAGGACAGGTGGCGGAGGCGCGCGCAAG 840
OY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaAlaPheGlnAlaArgLeuLys 280
DB 841 CTGGAGAGACAGCGCCAGCATACCTGACAGCGCGAGGCGCTCCAGGCGCGCGCTCAG 900
OY 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 901 ACCTGCTTCGAGCCCTGTGTGGAAGACATGACGCGCGCATGTGGCGCGGTGTGGAGAG 960
OY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
```

```
DB 961 GTTCAGAGCTGCGCTGGGACACAGCGCGCGCTGTGCTCCAGCAGCATACAC 1011
RESULT 8
US-09-954-456-760
; Sequence 760, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 760
; LENGTH: 1157
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-760
Alignment Scores:
Pred. No.: 2.43e-132 Length: 1157
Score: 1572.00 Matches: 314
Percent Similarity: 99.37% Conservative: 1
Best Local Similarity: 99.05% Mismatches: 2
Query Match: 99.18% Indels: 0
DB: 10 Gaps: 0
US-09-827-854-19 (1-317) x US-09-954-456-760 (1-1157)
OY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 62 ATCAAGGTTCTGTGGCTGCTGTGTCACATTCCTGACAGATGCCAGGCAAGGTG 121
OY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrPgiInser 40
DB 122 GAGCAAGCGGTGAGACAGACGCGAGCGCGAGCTCGCCAGCAGCGAGTGGCAGAGC 181
OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 182 GGGCAGCGCTGGGAAGTGGACACTGGCTGCTTTGGGATTACCTGGCGTGGTGACAGACA 241
OY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 242 CTGTCTGACAGACAGTGCAGAGAGCTGCTCAGCTCCCAAGTCAACCAGAACTGAGGGCG 301
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu 100
DB 302 CTGATGACAGACCATGAAAGAGTGAAGGCGCTACAAATCGGAATGGAGGAAACACTG 361
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
```


|||||
Db ACCCCGGTACGAGAGAGAGCGGCGCAGCTGCCAAGAGCTGCAGCGCGCGAGCC 421
QY 121 ATGLeuGluAlaAspMetLysPValCysGlyArgLeuValGlnTyrArgGlyVal 140
Db 422 CCGCTGGGCGGACATGGAGAGCGTGTCCGCCCTGTGTGTGACAGACCGCGGAGGTG 481
QY 141 GlnAlaMetLeuGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCATCTCGGCGAGAGCAGCGAGCGTGGGTCCGCCCTCCCACTCCGCC 541
QY 161 LysLeuArgGlnArgLeuLeuArgPspAlaAspPLeuGlnLysArgLeuAlaValTyr 180
Db 542 AAGCTGCGTAAAGCGGCTCTCCGCGATCCCGATGACCTCAGAAAGCGCTGCAGAGTAC 601
QY 181 GlnAlaGlyAlaArgGluValArgValArgValLysSerValIleArgGluArgLeuGly 200
Db 602 CAGGCGGGGGCCCGCGAGGCGCGCGAGCGGCTTCAGCGCATCCGCGAGCGCTGGGG 661
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
Db 662 CCCCTGTGTGAACAGGCGCGCGTGGCGGCGCAGCTGTGGCTCCCTGGCGCGCGAGCG 721
QY 221 LeuGlnGluArgAlaGlnAlaThrPglLysArgLeuArgAlaArgMetGluMetGly 240
Db 722 CTACAGAGAGCGGCGCGAGCGCTGGCGCGAGCGGCTGCGCGCGATGAGAGAGATGGCG 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 782 AGTCGAGACCGGCGAGCGCTGTGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 841
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 842 CCGGAGGAGACAGCGCGAGAGATACGCCCTGCGAGCGCGAGAGCTTCAGAGCGCGCTCAAG 901
QY 281 SerTrpPheGluProLeuValGlnLysPMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 902 ACCTGTTGACGCCCTGTGTGAGAGACATGCAGCGCAGTGGCGCGCTGTGTGGAGAG 961
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGCAAGGCTGCGGTGGGACACAGCGCGCTGTGTGCCAGGACATATAC 1012

RESULT 9
US-09-880-107-2244
; Sequence 2244, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Iwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2244
; LENGTH: 1157
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M12529
US-09-880-107-2244

Alignment Scores: 2,43e-132 Length: 1157
Pred. No.: 1572.00 Matches: 314

Percent Similarity: 99.37% Conservative: 1
Best Local Similarity: 99.05% Mismatches: 2
Query Match: 99.18% Indels: 0
DB: 10 Gaps: 0

US-09-827-854-19 (1-317) x US-09-880-107-2244 (1-1157)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAGTTCTGTGGGCTCGCTTGTGTCACATTCCTGGCAGAGTCCAGCGCAAGGTG 121
QY 21 GlnGlnAlaValGluThrLysProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 122 GAGCAAGCGGTGGAGACAGACGCGAGCGCGAGCTGCCAGACAGACAGAGTGGCAGAGC 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 182 GGCACAGCTGGGAACTGACACTGGGTGCTTTTGGATTACCTGCGGTGGTGCAGAGA 241
QY 61 LeuSerGluValGlnGluLeuLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 242 CTGTCTGACAGAGTGCAGAGAGAGCTGTACGTCCCAAGTCAACCAAGACTGAGAGGGG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluLeu 100
Db 302 CTGATGAGACAGACCATGAAGAGATTGAAGGCTTCAAAATCGGAATCGAGAGAACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 362 ACCCGGTGAGGAGAGAGAGCGGCGCGCGCTGTCCAAAGAGCTGCAGAGCGCGAGGCC 421
QY 121 ArgLeuGlyAlaAspMetLysPValCysGlyArgLeuValGlnTyrArgGlyVal 140
Db 422 CCGCTGGGCGGACATGGAGAGCGTGTGCGCGCTGTGTGACAGTACCGCGCGAGGTG 481
QY 141 GlnAlaMetLeuGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCATCTCGGCGAGAGCAGCGAGCGAGCTGGGTGCGCTCCCTCCCACTCGCGCG 541
QY 161 LysLeuArgGlnArgLeuLeuArgPspAlaAspPLeuGlnLysArgLeuAlaValTyr 180
Db 542 AAGCTGCGTAAAGCGGCTCTCCGCGATCCCGATGACCTCAGAAAGCGCTGCAGAGTAC 601
QY 181 GlnAlaGlyAlaArgGluValArgValArgValLysSerValIleArgGluArgLeuGly 200
Db 602 CAGGCGGGGGCCCGCGAGGCGCGCGAGCGGCTTCAGCGCATCCGCGAGCGCTGGGG 661
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
Db 662 CCCCTGTGTGAACAGGCGCGCGTGGCGGCGCAGCTGTGGCTCCCTGGCGCGCGAGCG 721
QY 221 LeuGlnGluArgAlaGlnAlaThrPglLysArgLeuArgAlaArgMetGluMetGly 240
Db 722 CTACAGAGAGCGGCGCGAGCGCTGGGCGAGCGGCTGCGCGCGATGAGAGAGATGGCG 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 782 AGTCGAGACCGGCGAGCGCTGTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 841
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 842 CCGGAGGAGACAGCGCGAGAGATACGCCCTGCGAGCGCGAGAGCTTCAGAGCGCGCTCAAG 901
QY 281 SerTrpPheGluProLeuValGlnLysPMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 902 ACCTGTTGACGCCCTGTGTGAGAGACATGCAGCGCAGTGGCGCGCTGTGTGGAGAG 961
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGCAAGGCTGCGGTGGGACACAGCGCGCTGTGTGCCAGGACATATAC 1012

RESULT 10
US-09-827-854-7

Qy	158	sleuArgrgysleuAaVgGInIarXleuLeuAhrqspAlaAspApleuGInIlyArqLeuA1	178
Db	21336	CCTGGCGAAGCTGyGTAAAGCGGCTCTCCCGCATGCACTGACAGAAAGTGGCTGGC	21395
Qy	178	aValTrgInaIaGlaYlaAaArgUglYlaAGlaGuaArgGlyleuSerAaIleArqGluAr	198
Db	21396	AGTGTACCAAGCGCGGGGCGCCGAGAGGGCGCCGAGCGCGGCTTCACAGCCATCCGAGACG	21455
Qy	198	gIeuNlyrProleuVaIglUngInIlyArqValArqAlaArpVaIaArpVaIglYserLeuAlaGl1	218
Db	21456	CCTGGGGCCCCCTGGGTGGAACAGAGGCGCGTGGGGCGCGCCACTGTGGGTCTCCTTGGCCGG	21515
Qy	218	YglInrProleuGInIuaArgAlaGlaAlaATrpbIyglUaArqLeuArqAlaArqMetGlu1	238
Db	21516	CCAGCGCGTACAGAGAGCGGGCCCGACGCTGGGGCGAGCGCGCTGGCCGCGCGGATGGAGA	21575
Qy	238	uMetGlyserArqThrArqAspArqLeuAspGluValYsgIuGInVaIAlaGluValAr	258
Db	21576	GATGGCGACCGCGAGCCCGGACCGCTGGAGACAGGTGAAGAGACAGATGGCGGAGGTGGC	21635
Qy	258	gAlaIyrlseuGluGluGInIaAGInIrlleArqleuGInaIaGlaAhpAGIAlaAr	278
Db	21636	CGCCAAAGCTGGAGAGAGAGGCCACAGATTCAGCTGCAGGGCGAGGCTTCACAGGCCG	21695
Qy	278	gIeuYsSerTrpPheGluPProleuVaIglUaSpMetGInaArqGInTrpAlaGlyLeuVa	298
Db	21696	CCTCAAGAGCTGTGTCGAGCCCTGTGTGAACAACATGCAGGCCCAATGGGCGGGCTGCT	21755
Qy	298	IglUlyrVaIglAlaIaIaVaIglYThrSerAlaIaProValProSerAspAsnHis	317
Db	21756	GGAGAAAGGTGCAGGCTGTGCCGTGGGACCAAGCGCCGCCCTGTGTGCCAGGACATTCAC	21813
RESULT 13			
US-09-925-302-133			
Sequence 133: Application US/09925302			
Patent No. US20020044941A1			
GENERAL INFORMATION:			
APPLICANT: Rosen et al.			
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies			
FILE REFERENCE: PA104			
CURRENT APPLICATION NUMBER: US/09/925,302			
CURRENT FILING DATE: 2001-08-10			
PRIOR APPLICATION NUMBER: PCT/US00/05918			
PRIOR FILING DATE: 2000-03-08			
PRIOR APPLICATION NUMBER: 60/124,270			
PRIOR FILING DATE: 1999-03-12			
NUMBER OF SEQ ID NOS: 896			
SOFTWARE: PatentIn Ver. 2.0			
SEQ ID NO 133			
LENGTH: 786			
TYPE: DNA			
ORGANISM: Homo sapiens			
US-09-925-302-133			

Alignment Scores:				
Pred. No.:	1,79e-69	length:	786	
Score:	867.50	Matches:	183	
Percent Similarity:	98.39%	Conservative:	0	
Best Local Similarity:	98.39%	Mismatches:	3	
Query Match:	54.73%	Indels:	2	
DB:	10	Caps:	0	
US-09-827-854-19 (1-317) x US-09-925-302-133 (1-786)				
OY	132	ArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThrGlnGluLeu	151	
Db	6	CGCCTGTGCAGTACCGGGGAGAGTGCASGSCCATGTCTGGCCACGACCGAGAGACTG	65	
OY	152	ArgValArgLeuAlaSerHisLeuArgLysLeuArgGlnArgLeuLeuArgAspAlaAsp	171	
Db	66	CGGGGCGCCTCGCCTCCACCTCGGCA-CTGGTAA-CGGCTCTTCGCGATGCCAT	123	

Oy	172	AspleuglnlvsarvgleuvalaValTTCglnalaglialarvglnugllyalaglIuargly	191
Db	124	GACCGCAAGAGCCCTCGGACGTACACAGCCGGGGCCCGCAGAGGCCCGCAAGCGGGC	183
Oy	192	LeuSerAlaIlearggluIuargLeuglyProLeuValIgluInglyIarvAlarvAlaIa	211
Db	184	CTCAGCGCCATCCCGAGAGCGCTGGGGGCCCTGGTGGAAACAGGGCCCGCTGGCGGGCGCC	243
Oy	212	ThrValIgySerLeuAlaIglvInProLeuInglnIuargIaIaIarvpglygluIarg	231
Db	244	ACTGTGGGCTCCCTCGGCCCGGCCACCGCTACAGAGAGCGGGCCAGGGCTTGGGGCGAGCGG	303
Oy	232	LeuArgAlaArvMeIugluIuMeGlySerarvThraIrgasparIeuaSpgluValIys	251
Db	304	CTGGCGCGCGGATGGAGAGATGGGACCGCGACCCGCGACCGCTCGACGAGAGTGTAAG	363
Oy	252	GluglnValAlaIgluValArgAlaIlylsIeuglnIuglnIaIaInglnIleargIeugln	271
Db	364	GAGCAGGTGGGGAGAGTCCGGCCAAACCTGGAGGAGAGAGGCCCGCAAGATACCGCTTCAG	423
Oy	272	AlaIguAlaIpheglnAlaIargLeuIysSerTrpIhegluProLeuValIguAspMeGln	291
Db	424	GCCGAGGECCTCCAGAGCCCGCGCTCAAGAGCTGGTTCAGAGCCCGCTGGTGGAAACATTCAG	483
Oy	292	ArgcIInrrpalagIlyleuValIgluIysValIcIlnAlaIaValIglyThrSerAlaIaPro	311
Db	484	CGCCAGTGGCGCGGGCTGGTGGAGAAAGTGCAGGCTCCGTGGGCAACACGCCGCCCT	543
Oy	312	ValProSerAspAsnHis	317
Db	544	GTGGCCAGCGCAATTCAC	561

```

RESULT 14
US-09-964-824A-374/c
: Sequence 374, Application US/09964824A
: Patent No. US20020102531A1
:
: GENERAL INFORMATION:
:   APPLICANT: Horrigan, Stephen
:   TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
:   TITLE OF INVENTION: Sets
:   FILE REFERENCE: 689290-73
:   CURRENT APPLICATION NUMBER: US/09/964, 824A
:   CURRENT FILING DATE: 2001-09-27
:   PRIOR APPLICATION NUMBER: US/60/236,033
:   PRIOR FILING DATE: 2000-09-28
:   PRIOR APPLICATION NUMBER: US/60/236,032
:   PRIOR FILING DATE: 2000-09-28
:   PRIOR APPLICATION NUMBER: US/60/236,028
:   PRIOR FILING DATE: 2000-09-28
:   NUMBER OF SEQ ID NOS: 583
:   SOFTWARE: PatentIn version 3.0
:   SEQ ID NO 374
:   LENGTH: 478
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: FEATURE:
:   NAME/KEY: misc_feature
:   LOCATION: (1)...(478)
:   OTHER INFORMATION: n=a,t,g or c
US-09-964-824A-374

```

Alignment Scores:		
Pred. No.:	5, 066-48	Length: 478
Score:	625.00	Matches: 126
Percent Similarity:	97.63%	Conservative: 1
Best Local Similarity:	96.92%	Mismatches: 2
Query Match:	39.43%	Indels: 1
DB:	10	Gaps: 0
US-09-827-854-19 (1-317) x US-09-964-824A-374 (1-478)		
0y	1 MetcysValLeuTrrPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20y

Db 396 ATGAAGGTTCTGTGGGCTGCTGCTGTGCACATTCTGGCAGATGCCAGGCCAAGGTG 337
Qy 21 GUGlnAlaValaGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 336 GAGCAAGCGGTGGAGACAGAGCCGAGCCNCAGCTGCCGACAGACCAGAGTGGCAGAGC 277
Qy 41 GUGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValaGlnThr 60
Db 276 GGCCAGCGCTGGAGACTGACACTGCTGCTTTTGGGATTTACTCTGGCTGGGTGCAGACA 217
Qy 61 LeuSerGluGlnValaGlnGluLeuLeu-SerSerGlnValaTrnGlnGluLeuArgAl 80
Db 216 CTGCTGACAGCTGACAGAGAGCTGCTCGAGCTCCAGGTCCACCCAGGAAGTGAAGGC 157
Qy 80 aleuMetAspGluTrpMetLeuGlyAlaTyrTrpSerGluLeuGlnGluGlnLe 100
Db 156 GCTGATGACGAGACCATGAAAGAGTTGAAAGCTTACAAATGGAATGGAGAACAACT 97
Qy 100 uThrProValaAlaGluGluThrArgAlaArgLeuSerGluLeuGlnAlaAlaGlnAl 120
Db 96 GACCCCGGTGGCGAGAGAGACCGGGCAGCGCTGTCCAAGAGCTGCAAGGGCGCGAGAGN 37
Qy 120 aArgLeuGlyAlaAspMetGluAspVal 129
Db 36 CCGGCTGGGCGGAGACATGGAGACGCTG 9

RESULT 15

US-09-880-107-2491/c
: Sequence 2491, Application US/09880107
: Patent No. US20020142981A1
: GENERAL INFORMATION:
: APPLICANT: Horne, Darci T.
: APPLICANT: Vockley, Joseph G.
: APPLICANT: Scherf, Uwe
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
: FILE REFERENCE: 44921-5028-WO
: CURRENT APPLICATION NUMBER: US/09/880,107
: CURRENT FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: PRIOR FILING DATE: 2000-10-02
: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2491
: LENGTH: 478
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20020142981A1 N33009
: NAME/KEY: unsure
: LOCATION: (1)..(478)
: OTHER INFORMATION: n = a or c or g or t
US-09-880-107-2491

Alignment Scores:

Pred. No.:	5,06e-48	Length:	478
Score:	625.00	Matches:	126
Percent Similarity:	97.69%	Conservative:	1
Best Local Similarity:	96.92%	Mismatches:	2
Query Match:	39.43%	Indels:	1
DB:	10	Gaps:	0

US-09-827-854-19 (1-317) x US-09-880-107-2491 (1-478)

Qy 1 MetLysValLeuTrpAlaAlaLeuValaThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 396 ATGAAGTTCGTGTGGCTGCTGCTGTGCACATTCTGGCAGATGCCAGGCCAAGGTG 337
Qy 21 GUGlnAlaValaGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 336 GAGCAAGCGGTGGAGACAGAGCCGAGCCNCAGCTGCCGACAGACCAGAGTGGCAGAGC 277

Qy 41 GUGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValaGlnThr 60
Db 276 GGCCAGCGCTGGAGACTGACACTGCTGCTTTTGGGATTTACTCTGGCTGGGTGCAGACA 217
Qy 61 LeuSerGluGlnValaGlnGluLeuLeu-SerSerGlnValaTrnGlnGluLeuArgAl 80
Db 216 CTGCTGACAGCTGACAGAGAGCTGCTCGAGCTCCAGGTCCACCCAGGAAGTGAAGGC 157
Qy 80 aleuMetAspGluTrpMetLeuGlyAlaTyrTrpSerGluLeuGlnGluGlnLe 100
Db 156 GCTGATGACGAGACCATGAAAGAGTTGAAAGCTTACAAATGGAATGGAGAACAACT 97
Qy 100 uThrProValaAlaGluGluThrArgAlaArgLeuSerGluLeuGlnAlaAlaGlnAl 120
Db 96 GACCCCGGTGGCGAGAGAGACCGGGCAGCGCTGTCCAAGAGCTGCAAGGGCGCGAGAGN 37
Qy 120 aArgLeuGlyAlaAspMetGluAspVal 129
Db 36 CCGGCTGGGCGGAGACATGGAGACGCTG 9

Search completed: March 15, 2003, 03:05:08
Job time : 80.8753 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:09:47 ; Search time 1289.69 Seconds
(without alignments)

3980.771 Million cell updates/sec

Title: US-09-827-854-19

Perfect score: 1585

Sequence: 1 MKVLMALVTFGLGCAKV.....VEKVOAAGTSAPPSDNH 317

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_p2n.model -DEV=xlp
-O=/cgn2.1/USPTO/spool/US09827854/rtnat_11032003_101610_27495/app_query.fasta.1.3576
-DB=EST -QFMT=fastap -SUFMT=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854_qcgn.1.1.8826_ernat_11032003_101610_27495 -NCPU=6 -ICPU=3
-NO_XLPTX -NO_MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlm:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1286.5	81.2	1027	13	BI670367
2	1250	78.9	842	12	BG763371
3	1244	78.5	933	13	BI668318
4	1216	76.7	938	12	BG761746
5	1213.5	76.6	942	13	BI600906
6	1199	75.6	800	13	BM042094
7	1198.5	75.6	927	12	BG472299
8	1195.5	75.4	922	13	BI597743
9	1186.5	74.9	817	12	BG774871
10	1185	74.8	811	13	BI600563
11	1171.5	73.9	845	12	BG829472
12	1148	72.4	790	12	BG707147
13	1145	72.2	919	13	BI551475
14	1135	71.6	757	13	BM042228
15	1125.5	71.0	907	12	BG706129
16	1124	70.9	706	14	BM728696
17	1124	70.9	741	12	BG762924
18	1122	70.8	1110	11	AK010261
19	1120	70.7	804	12	BG702752
20	1119	70.6	834	13	BM042676
21	1116.5	70.4	808	13	BI668329
22	1113	70.2	855	13	BI616362
23	1110	70.0	803	13	BI670350
24	1107	69.8	782	12	BG716776
25	1107	69.8	812	13	BI601551
26	1107	69.8	965	14	BQ677266
27	1104	69.7	757	13	BI603658
28	1101.5	69.5	797	12	BG715366
29	1101	69.5	846	13	BI159757
30	1100	69.4	796	13	BM042153
31	1095	69.1	757	12	BG707750
32	1095	69.1	794	12	BI601279
33	1095	69.1	954	12	BF967543
34	1093	69.0	790	13	BI551066
35	1092	68.9	790	13	BI551811
36	1088.5	68.7	798	12	BG708414
37	1087.5	68.6	891	13	BI548292
38	1081.5	68.2	802	13	BI458355
39	1080	68.1	1100	14	BM914382
40	1079.5	68.1	914	13	BI603523
41	1079	68.1	748	13	BI553085
42	1078.5	68.0	930	12	BF967857
43	1074.5	67.8	812	12	BG769968
44	1074	67.8	688	13	BI544886
45	1073.5	67.7	803	12	BG709360

ALIGNMENTS

RESULT 1
BI670367
LOCUS 603292738P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5312024 5',
DEFINITION B1670367.1 GI:15584600
ACCESSION B1670367
VERSION B1670367.1
KEYWORDS mRNA sequence.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1027)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

195 GGCCAGCGCTGGGAACTGGCTGGCTGCTTTGGATTACCTGGCTGGTGACAGACA 254
61 LeuSerGluGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuAArgAla 80
255 CTGTCTGACAGAGTGCAGAGAGAGCTGCTCAGCTCCAGGTACCCAGAGACTGAGAGGGC 314
81 LeuMetAspGluThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnGlnLeu 100
315 CTGATGAGCAGACCATGTAAGAGAGTTGAAGGCTTACAAATCGGAACCTGAGAGAACACTG 374
101 ThrProValAlaGlnGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
375 ACCCGGCTGGCGAGAGAGACGGCGGACGCTGTCCAAAGAGCTGCAGCGCGCGAGGCC 434
121 ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
435 CGGCTGGGGCGGACATGAGAGAGCTGTGCGCGCTGGTGTGACGACCGCGCGAGGAG 494
141 GlnAlaMetLeuGlnGlnSerThrGlnGlnGlnValArgLeuAlaSerHisLeuArg 160
495 CAGGCCATCTCGCGCAGAGACACCGAGAGCTGCGGGTGGCTCCCTCCACCTGCCG 554
161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
555 AAGCTGCGTAAGCGGCTCTCCGCGATGCCGATGACCTGCAGAGCGCTGGCAGTGTAC 614
181 GlnAlaGlnAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
615 CAGCGCGGGGGCGCGAGGCGCGCGAGCGCGCTCAGCGCCATCGCGAGCGCGCTGGGG 674
201 ProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
675 CCCCTGCTGAACAGAGCGCGCTGCGCGCGCGCACGTGGGGCTCCGCGCGCGCGAGCG 734
221 LeuGlnGlnAlaArgAlaGlnAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 239
735 CTACACAGACCGGCGCGCGCTGGCGGAGCGCTGCGCGCGGAGGAGGACACATTG 794
239 TcLysSerArgThrArgAspArgLeuAsp-GluValLysGlnGlnGlnGlnGlnGlnGlnGln 259
795 GGGCAGCGCGGACCGCGGAGCGCTGGAAGAGGTGAAGAGAGCGGCGGAACTTTCGCCG 854
259 LbLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 279
855 GCCAAGGCTGGGGAGACCGAGCGCGCAATTCGCTTGCAAGGCCAA-GCCTTCCAGCGCGCTG 913
279 eulLysSerTrrpPheGlu 284
914 CCAAGAGCTGCTTCAAA 930

RESULT 4
BG761746 938 bp mRNA linear EST 15-MAY-2001
LOCUS BG761746
DEFINITION 602117942F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4841411 5',
mRNA sequence.
ACCESSION BG761746
VERSION BG761746.1 GI:14072399
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 938)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: ATCC/DC/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L16M1674 row: c column: 12
High quality sequence stop: 767.
Location/Qualifiers
1.938
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4841411"
/clone_1lb="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA Synthesis Kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

BASE COUNT 171 a 272 c 374 g 121 t
ORIGIN

Alignment Scores:
Pred. No.: 1.84e-117 Length: 938
Score: 1216.00 Matches: 271
Percent Similarity: 92.62% Conservative: 5
Best Local Similarity: 90.94% Mismatches: 11
Query Match: 76.72% Indels: 11
DB: 12 Gaps: 2

US-09-827-854-19 (1-317) x BG761746 (1-938)

QY 1 MetLysValLeuThrPalaLeuLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
DB 57 ATCAAGCTTCTGGGGCTCGTGGTGGTCACTTCTCTGCGAGAGTCCAGGCCAAGGTTG 116
QY 21 GlnGlnAlaValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 40
DB 117 GAGCAAGCGGTGGAGAGACAGAGCGGAGCGCGAGCTGCGCAGAGACGAGTGGCAGAC 176
QY 41 GlyGlnArgTrrpGlnLeuLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 60
DB 177 GGCCACGCTGGGAACTGCGACACTGGGCTCTTTGGATTACCTGCGCTGGGCGCAGACA 236
QY 61 LeuSerGluGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
DB 237 CTGTCTGAGCAGGTGCAGAGAGAGCTGTCTCAGCTCCAGGTACCCAGAGACTGAGGGCG 296
QY 81 LeuMetAspGluThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnGlnLeu 100
DB 297 CTGATGAGCAGACCATGTAAGAGAGTTGAAGGCTTACAAATTCGGAACCTGAGAGAACACTG 356
QY 101 ThrProValAlaGlnGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
DB 357 ACCCGGCTGGCGAGAGAGACGGCGGACGCTGTCCAAAGAGCTGCAGCGCGCGAGGCC 416
QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
DB 417 CGGCTGGGGCGGACATGAGAGAGCTGTGCGCGCTGGTGTGACGACCGCGCGAGGAG 476
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
DB 477 CAGGCCATCTCGCGCAGAGACCGAGAGAGTGGGGTGGCTGCCCTCCACCTGCGCGC 536
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 537 AAGCTGCGTAAGCGGCTCTCCGCGATGCCGATGACCTCAGAGAGCGCTGGCAGTGTAC 596
QY 181 GlnAlaGlnAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 597 CAGGCCGGGGCGCGAGGCGCGCGAGCTGTCTCAGCGCCATTCGCGAGCGCTGGGG 656

QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySer-LeuAlaGlyGlnPro 220
 |||||
 Db 657 CCCCTGGTGGAAACAGGCGCGCTGCGGCGCCACACTGTGGCTCCCTTGGCGCGCCAGCC 716
 QY 220 OLeuGlnGluArgAlaGlnAlaTrp-GlyGluArg-LeuArgAlaArgMetGlu---Glu 238
 |||||
 Db 717 GCTACAGAGAGCGGCGCCAGCGCTGCGGCGAGCGCTGCGCGCGGATGAGAGACAT 776
 QY 239 MetGlySerArgThr-Arg-AspArgLeuAspGluValIys-GluGlnValAla---Glu 256
 |||||
 Db 777 TGGGACGCCGAGACCCCGGAGACCGCTTGACAGAGTGAAGGAGCAGGTGGCGGAPACG 836
 QY 257 ValArgAlaIysLeuGlnGlu-GlnAlaGlnGlnIleArgLeuGlnAla--GlnAlaPhe 275
 |||||
 Db 837 TTGGCGGCCACAGCTGAGAGACAGCGCCAGAGATACCCCTGACAGCGCGGAAAGGCTTC 896
 QY 276 GlnAlaArgLeuIysSerTrpPheGluProLeuVal 287
 |||||
 Db 897 CCAGCGCGGCTCAAAAGCTGTGCAAGCCCTGGTT 932

RESULT 5
 B1600906 942 bp mRNA linear EST 07-SEP-2001
 LOCUS 603249241P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5301010 5',
 DEFINITION mRNA sequence.
 ACCESSION B1600906
 VERSION B1600906.1 GI:15493845
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 942)
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@b-remail.nih.gov
 Tissue Procurement: Miklos Palikovit, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshitaki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 Place: LLM11762 row: a column: 11
 High quality sequence stop: 762.
 Location/Qualifiers
 source 1..942
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5301010"
 /clone_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescript (modified
 pBluescript KS+); Site:1: BamHI, Site:2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTT-3',
 size-selected for average insert size 2.3 kb and
 normalized to R0.5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NHGRI/NHGRI, National
 Institutes of Health). Note: this is a NIH-MGC Library."

BASE COUNT 177 a 284 c 361 g 120 t
 ORIGIN

Alignment Scores: 3.39e-117 length: 942
 Pred. No.: 1213.50 matches: 265
 Score: 93.47% Conservative: 7
 Percent Similarity:

Best Local Similarity: 91.07% Mismatches: 13
 Query Match: 76.56% Indels: 8
 DB: 13 Gaps: 1
 US-09-827-854-19 (1-317) x B1600906 (1-942)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyGlnAlaIysVal 20
 |||||
 Db 73 ATGAAAGTTCTGTGGGCTGCTGCTGTGTCACATTCCTGCGAGATGCCAGGCAAGTG 132
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnIleThrGluTrpGlnSer 40
 |||||
 Db 133 GAGCAACCGGTGGAGACAGACCGGAGCCGAGCTCCGACAGACAGACAGACAGAGTGGCAGAC 192
 QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGluTrp 60
 |||||
 Db 193 GGCACGCTGGGAACCTGACCTGGGTCTCTTTGGATTACCTGCGCTGGGTGACAGAC 252
 QY 61 LeuSerGluGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 |||||
 Db 253 CTGTCTGACAGAGTGCAGAGAGAGCTGCTCAGCTCCAGGTACCAAGAACTGAGGCGG 312
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu 100
 |||||
 Db 313 CTGATGACAGAGACCATGAAAGAGTTGAAGGCTTACAAATCGGAACCTGAGAGAACTG 372
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGlnGlnAlaAlaGlnAla 120
 |||||
 Db 373 ACCCGCTGGCGAGAGAGCGCGGCGGCTGTCCAGAGCTGAGGCGGCGAGGCC 432
 QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
 |||||
 Db 433 CGGCTGGGCGGACATGAGAGACGTGTGCGGCTGTGTGAGTACCGCGGCGAGGTG 492
 QY 141 GlnAlaMetLeuGlnGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||
 Db 493 CAGGCGATCTGCGGCGAGACACCGAGAGCTGCGGCTCCCTCCACCTCGCGC 552
 QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||||
 Db 553 AAGCTCGTAAAGGGCTCTCCGCGATGCGGATGCGAGAGCTCGAAGGCGCTGAGAGTGTAC 612
 QY 181 GlnAlaGlnAlaArgGlnGlnAlaGlnGlyLeuSerAlaIleArgGlnArgLeuGly 200
 |||||
 Db 613 CAGCGCGGGGCGCGGAGGGCGCGGCGGCTGAGGCGCATCGCGAGCGGCTCGGG 672
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 |||||
 Db 673 CCCCTGGTGGAAACAGGCGCGGCGGCGGCGGCTGAGGCGCGGCGGCGGCGG 731
 QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGlyArgLeuArgAla-ArgMetGlu-GluMetG 240
 |||||
 Db 732 CTACAGAGAGCGGCGCGGCTGCGGCGGAGCGGCTGCGGCGGATGAGGAGCAATGG 791
 QY 240 LysSerArgThrArgAspArgLeuAsp-GluValIys-GluGlnValAlaGluVal-Arg 259
 |||||
 Db 792 GAGCGCGGAGCCGCGAGCGGCTGAGCGAGGAGCAAGGTGGGAAAGTTCGCGG 851
 QY 259 LysLeuGlnGluGlnAla---GlnGlnIleArgLeuGlnAlaGluAlaPheGlnAla 278
 |||||
 Db 852 CCACACTTGGAGAGACACGCGCCAGGCAAGATACGCTTGCAGGCGGAGGCTTCACGCC 911
 QY 278 rGluLeuSerTrpPheGluPro 285
 |||||
 Db 912 GCCTCAAAAC-TGGTTGACCCCT 933

RESULT 6
 BM042094 800 bp mRNA linear EST 07-NOV-2001
 LOCUS BM042094
 DEFINITION 603615713P1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5420618 5',
 mRNA sequence.
 ACCESSION BM042094
 VERSION BM042094.1 GI:16771361

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 800)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE Unpublished (1999)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTP/DRP

cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM1875 row: 1 column: 03
High quality sequence stop: 792.
Location/Qualifiers

FEATURES
source 1..800

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5420618"
/clone_lib="NIH_MGC_112"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOMB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."

BASE COUNT 146 a 239 c 315 g 100 t
ORIGIN

Alignment Scores:

Pred. No.: 8.93e-116 length: 800
Score: 1199.00 Matches: 243
Percent Similarity: 99.20% Conservative: 5
Best Local Similarity: 97.20% Mismatches: 2
Query Match: 75.65% Indels: 1
DB: 13 Gaps: 0

US-09-827-854-19 (1-317) x BM042094 (1-800)

QY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
DB 50 ATGAAGATTCTGTGGCTGGCTGTGTCACATTCCTGGCAGATGCCAGGCAAGGTG 109
QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnProGlnSer 40
|||||
DB 110 GAGCAAGCGGTGGAGACAGCGCGAGCCGACCTCGCCAGAGACCGAGTGGCAGAGC 159
QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
DB 170 GGCACAGCGTGGAACTGGCAGCTGGCTTTGGGATTACTGGCTGGGTGGCAGACA 229
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
|||||
DB 230 CTGTGTGACAGAGTGCAGAGGAGTGTCTCAGCTCCAGGTACCCAGGAATGAGGGCG 289
QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
|||||
DB 290 CTGATGGACAGACCATGAGAGGTGAAGGCTTACAAATCGGAATGGAGGAAACAATG 349
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
|||||
DB 350 ACCCGGTTGGCGAGAGACGCGGCGACGGCTCTCCAGAGAGCTGCAAGCGGCGCAGGCC 409

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
DB 410 CGCGTGGCGCCGACATGAGGAGACGTGTGGCGCGCTGTGTGACATGCCGCGAGGTG 469
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
DB 470 CAGGCCATGCTCGGCCAGAGACACCGAGAGACTGCGGGTGGCTTCCCTCCACTTGGC 529
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
DB 530 AACCTGGTGAAGGCGCTCCCGCGATGCCGATGACCTGCAGAAAGCGCTGGCAGGTAC 589
QY 181 GlnAlaGlyAlaArgGlnGlyAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||
DB 590 CAGGCGGGGCGCCGCGAGGCGCGGAGCGCGCTCAGGCCATCCGCGAGCGCTGGGG 649
QY 201 ProLeuValGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
|||||
DB 650 CCGCTGGTGGACAGGGCGCGTGGCGGCCAGCACTGTGGCTCCCTGGCGGCCAGCCG 709
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGlnMetGly 240
|||||
DB 710 CTACAGAGAGCGGCGCCAGGCGCTGGGGGAGCGGTG-CGCGCGCGATGAGAGATGGCG 768
QY 241 SerArgThrArgAspArgLeuAspGluVal 250
|||||
DB 769 ACCGCGGACCGCGACGCTGAGCAGAGTG 798

RESULT 7
BG472299 927 bp mRNA linear EST 21-MAR-2001
LOCUS
DEFINITION 602513830F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4645759 5',
mRNA sequence.
ACCESSION BG472299
VERSION BG472299.1 GI:13404485

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 927)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE Unpublished (1999)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM1419 row: k column: 08
High quality sequence stop: 848.
Location/Qualifiers

FEATURES
source 1..927

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4645759"
/clone_lib="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOMB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC Library."

BASE COUNT 170 a 271 c 371 g 115 t
ORIGIN

Alignment Scores:

Pred. No.:	1.25e-115	Length:	927
Score:	1198.50	Matches:	274
Percent Similarity:	92.33%	Conservative:	3
Best Local Similarity:	91.33%	Mismatches:	13
Query Match:	75.62%	Indels:	12
DB:	12	Gaps:	1

US-09-827-854-19 (1-317) x BG472299 (1-927)

```

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
  |||||
DB 47 ATGAAGTTCTGTGGCTCGTTGCTGTCACATTCTTGGCAGAGTCCAGGCCAAGGTG 106
QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnThrGlnPheGlnSer 40
  |||||
DB 107 GAGCAAGCGGTGGAGACACAGCCGAGCCGAGCTGCCAGACAGCCAGAGTGGCAGAC 166
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
  |||||
DB 167 GGCCAGCCCTGGGAACCTGCGACTGGGTCTTTGGATTACTGCGCTGGGTGGCAGACA 226
QY 61 LeuSerGluGlnValGlnGluLeuLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
  |||||
DB 227 CTGTCTGACAGGTGCAGAGAGAGAGAGCTCCTCCAGTCCAGTCCAGGACTGAGAGCGG 286
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnGlnLeu 100
  |||||
DB 287 CTGATGAGACAGACACCTGTAAGAGATTGAAGGCTTACAAATTCGAACTGAGAGACACTG 346
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
  |||||
DB 347 ACCCCGCTGGCGAGAGAGACGCGCGACGCTGTCCAAAGACTGCAGCGCGCGCAGCGCC 406
QY 121 ArgLeuGlyAlaAspMetLysPalaLysGlyArgLeuValGlnThrArgGlyGlnVal 140
  |||||
DB 407 CGGCTGGGGCGGACATGAGAGAGCTGCGCGCTGTGTGACAGTCCGCGCGAGGTG 466
QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
  |||||
DB 467 CAGGCGATCTCGGCGCAGAGACCCGAGGCTGCGGCTGCGCTCCCACTCCAGCTCGCC 526
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTrp 180
  |||||
DB 527 AAGCTGCGTAAAGCGGCTCTCCGCGATGCCATGACTCACAAGGCGCTGCGAGTCTAC 586
QY 181 Gln-AlaGlyAlaArgGluGlyAlaGlnArgGly--LeuSerAlaAlaLeuArgGlnLeu 199
  |||||
DB 587 CAGGGCGGGGGCGCGGAGGCGCGCGCCCTCAAGGCGCATCCGCGAGCGCCCTT 646
QY 200 --GlyProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyG 219
  |||||
DB 647 GGGGCGCCCTGGGTGGAACAGGCGCGGCTGCGGCGCCACTGTGGCTCTGCGCGGCC 706
QY 219 InProLeuGlnGluArgAlaGlnAlaAlaTrpGlyGluArgLeuArgAlaArgMetGlnGlu 239
  |||||
DB 707 AACG-CTACAGAGAGCGGCGCCAGGCTGGGGCGAAGCGTG-CGCGGGCGGATGAGAGAGA 764
QY 239 etGlySerArg-TrpArgAspArgLeuAspGluValLysGluGlnValAlaGlu-ValAr 258
  |||||
DB 765 TGGGGCAGCGGGAGCCCGGACCTGAGCGAGGTGAAGAGACAGTGGCGAGGCTGG 824
QY 258 GAlaLysLeuGlnGluGlnAlaGlnGlnLeuArgLeuGlnAlaGlnAla---PheGlnAl 277
  |||||
DB 825 CCGCAAGCTGGAGAGAGACAGCGCCAGCAGATACGGCTGCAGGCGCGAGGCGCTTCCAGGGC 884
QY 277 aArg-LeuLysSerTrpPhe-GluProLeuValGluAspMet 290
  |||||
DB 885 CCGGCTCAAGAGAGCTGTGCGAGCCCTGGGTGGAAGACATG 926
  |||||
RESUL 8
BIS97743 LOCUS BIS97743 922 bp mRNA linear EST 07-SEP-2001

```

```

DEFINITION 603248609F1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:5300259 5',
ACCESSION mRNA sequence.
VERSION BIS97743
KEYWORDS BIS97743.1 GI:15490682
SOURCE EST.
ORGANISM human.
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
  1 (bases 1 to 922)
  NIH-MGC http://mhc.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: c9apbs-remail.nih.gov
  Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
  cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
  Toshiyuki and Piero Carninci (RIKEN)
  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LAM11760 row: b column: 04
  High quality sequence stop: 782.
  Location/Qualifiers
    1..922
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="IMAGE:5300259"
      /clone.lib="NIH_MGC_96"
      /tissue.type="hypothalamus"
      /lab.host="DH10B"
      /note="Organ: brain; Vector: pBluescript (modified
      pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
      ); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
      size-selected for average insert size 2.3 kb and
      normalized to R07 5. This is a primary library enriched
      for full-length clones and constructed using the
      cap-trapper method (carninci, in preparation). Library
      constructed by M. Brownstein (NHGRI, National
      Institutes of Health). Note: this is a NIH_MGC Library."
    BASE COUNT 176 a 273 c 359 g 114 t
    ORIGIN
  Alignment Scores:
    Pred. No.: 2.56e-115 Length: 922
    Score: 1195.50 Matches: 256
    Percent Similarity: 93.55% Conservative: 5
    Best Local Similarity: 91.76% Mismatches: 14
    Query Match: 75.43% Indels: 5
    DB: 13 Gaps: 1
  US-09-827-854-19 (1-317) x BIS97743 (1-922)
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
  |||||
DB 73 ATGAAGTTCTGTGGCTCGTTGCTGTCACATTCTTGGCAGAGTCCAGGCCAAGGTG 132
QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnThrGlnPheGlnSer 40
  |||||
DB 133 GAGCAAGCGGTGGAGACACAGCCGAGCCGAGCTGCCAGACAGCCAGTGGCAGAC 192
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
  |||||
DB 193 GGCCAGCGCTGGGAACCTGCGACTGGGTCTTTGGATTACTGCGCTGGGTGGCAGACA 252
QY 61 LeuSerGluGlnValGlnGluLeuLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
  |||||
DB 253 CTGTCTGACAGGTGCAGAGAGAGAGCTGCTCAGCTCCAGTCCAGGACTGAGAGCGG 312
  |||||
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu 100
  |||||

```


Db	313	CGATGGACGACACATGAAGAGTTGAAGGCTTCAATATCGAATCGAGACGAACACTG	372	
OY	101	ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120	
Db	373	ACCCGCGTGGCGGAGAGACGCGGGCGACGGCTGTCCAAAGGACCTCAGCGCGCGACAGGC	432	
OY	121	ArgLeuGlyAlaAspMetGluAspValLysGlyArgLeuValGlnThrArgLysGluVal	140	
Db	433	CGGGTGGGGCGGACATGAGAGACTGTGGCGGCCCTGTGGTGCATATACCGCGCGAGGGT	492	
OY	141	GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg	160	
Db	493	CAGGCCATGCTGGCGCAGGACACGAGGAGCGTGGGGTGGCGCTGCCCTCCACCTGGGCG	552	
OY	161	LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180	
Db	553	AAGCTGCGTGAAGCGGCTCTCCGCGATGCGCGTACCTGCAGAAACCGCTGCACCTGTAC	612	
OY	181	GlnAlaGlyAlaArgGluGluGlyAlaArgLysArgGlyLeuSerAlaIleArgGluArgLeuGly	200	
Db	613	CAGGCCGGGGGCCCGGAGGGGCCCGACGCGGCTC-AGCGCCATCCGCGAGCGCTGGGG	671	
OY	201	ProLeuVal-GluGlnGlyArgValArgAlaIleThrValGlySerLeuAlaGlyLysL-P	220	
Db	672	CCCCGTGGTGGGAACAGGGCGGGCTCCGGCGCCGCACTGTGGGCTCCGTGGCGCGCAGGC	731	
OY	220	roLeuGlnGluArg-AlaGlnAlaIleArgLysGluArgLeuArgAlaArgMetGluMet	239	
Db	732	CGCTCAGAGAGCGGGGCCAGGCTGGGGCGGCGGATACGCGCGCATGACGACGACGAG	791	
OY	240	GlySerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAla	259	
Db	792	GGCAGCCGCGACCCGAGACCGCTGGACGAGTGAAGGAGACAGGTGGCGGGGTGGGGCCA	851	
OY	260	LysLeuGluGluGluGlnAlaGlnIleLeuArgLeuGlnAlaGluAlaPheGln	276	
Db	852	CAAGCTGAAGAAACACGCCCAAGAAATACGCTTA--GCCGAAGCGCTTCAGG	899	
RESULT 9	BG774871	817 bp	linear	EST 15-MAY-2001
LOCUS	BG774871			
DEFINITION	602649975p1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4760827 5',			
ACCESSION	BG774871			
VERSION	BG774871.1	GI:14045188		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
TITLE	1 (bases 1 to 817)			
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.			
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)			
	Unpublished (1999)			
	Contact: Robert Strausberg, Ph.D.			
	Email: cga@bbs-remail.nih.gov			
	Tissue Procurement: DCTD/DPF			
	cDNA Library Preparation: Ling Hong/Rudin Laboratory			
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)			
	DNA Sequencing by: Incyte Genomics, Inc.			
	Clone distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LNL at:			
	http://image.lnl.gov			
	Plate: L16CM1612	row: e	column: 20	
	High quality sequence stop: 813.			

FEATURES

- source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4760827"
/clone_lib="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"

```

Note:"Organ: prostate; Vector: pORF; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCCACGAG(C). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC library."

BASE COUNT	148 a	242 c	324 g	103 t
ORIGIN				

Alignment Scores:	
Pred. No.:	1,3e-114
Score:	1186.50
Percent Similarity:	97.32%
Best Local Similarity:	96.17%
Query Match:	74.86%
DB:	12
Length:	817
Matches:	251
Conservative:	3
Mismatches:	5
Indels:	0
Gaps:	4

US-09-827-854-19 (1-317) x BG774871 (1-817)

Oy	1	MettVsValIleuTrrPALalaleuIeuValIrrPheIeuIalagIcyGAlalalysal	20
Db	40	ATGAAGGTTCTGTGGCGTGGCGTTGCTGGTGCATTCCTGGCAGATGCCAGGCCAAGSTG	99
Oy	21	GIuGlnIalalValGIuThrGluProGluProGluIeuAArgIuGlnIuThrGluTrpIleuSer	40
Db	100	GAGCAAGCGGTGGACACAGAGCGGAGCGCCAGAGCTGGCCGACACACCAAGTGGCAGAGAC	150

QY 41 GLYGINARGTRPGIULUeAlaLeuGIYargPheTRpasPTYrLeuargGTRpValGlnThr 60

Db 160 GGCAGCGCTGGGAACCTGGCCTGGCTTTTGGGATTACCTGGCGCTGGGTGCAGACA 21:

61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80

Db 220 CTGCTGAGCAGGTGCAGGAGGAGCTGCTCAGCTCCAGGTCACCAGGAACTGAGGGC 2/

81 Leumetaspciuhrmetllysculdeulysalalyllysseruldeugucugmleu 100

DB 280 C1GAGAGAGACCA1GAAAGAG1AGAAAGCC1ACAAA1CGAAAC1AGAAAGAAACAAAC1G

Q	A
101	THE FIVE OF HEARTS
102	THE FIVE OF HEARTS
103	THE FIVE OF HEARTS
104	THE FIVE OF HEARTS
105	THE FIVE OF HEARTS
106	THE FIVE OF HEARTS
107	THE FIVE OF HEARTS
108	THE FIVE OF HEARTS
109	THE FIVE OF HEARTS
110	THE FIVE OF HEARTS
111	THE FIVE OF HEARTS
112	THE FIVE OF HEARTS
113	THE FIVE OF HEARTS
114	THE FIVE OF HEARTS
115	THE FIVE OF HEARTS
116	THE FIVE OF HEARTS
117	THE FIVE OF HEARTS
118	THE FIVE OF HEARTS
119	THE FIVE OF HEARTS
120	THE FIVE OF HEARTS
121	THE FIVE OF HEARTS
122	THE FIVE OF HEARTS
123	THE FIVE OF HEARTS
124	THE FIVE OF HEARTS
125	THE FIVE OF HEARTS
126	THE FIVE OF HEARTS
127	THE FIVE OF HEARTS
128	THE FIVE OF HEARTS
129	THE FIVE OF HEARTS
130	THE FIVE OF HEARTS
131	THE FIVE OF HEARTS
132	THE FIVE OF HEARTS
133	THE FIVE OF HEARTS
134	THE FIVE OF HEARTS
135	THE FIVE OF HEARTS
136	THE FIVE OF HEARTS
137	THE FIVE OF HEARTS
138	THE FIVE OF HEARTS
139	THE FIVE OF HEARTS
140	THE FIVE OF HEARTS
141	THE FIVE OF HEARTS
142	THE FIVE OF HEARTS
143	THE FIVE OF HEARTS
144	THE FIVE OF HEARTS
145	THE FIVE OF HEARTS
146	THE FIVE OF HEARTS
147	THE FIVE OF HEARTS
148	THE FIVE OF HEARTS
149	THE FIVE OF HEARTS
150	THE FIVE OF HEARTS
151	THE FIVE OF HEARTS
152	THE FIVE OF HEARTS
153	THE FIVE OF HEARTS
154	THE FIVE OF HEARTS
155	THE FIVE OF HEARTS
156	THE FIVE OF HEARTS
157	THE FIVE OF HEARTS
158	THE FIVE OF HEARTS
159	THE FIVE OF HEARTS
160	THE FIVE OF HEARTS
161	THE FIVE OF HEARTS
162	THE FIVE OF HEARTS
163	THE FIVE OF HEARTS
164	THE FIVE OF HEARTS
165	THE FIVE OF HEARTS
166	THE FIVE OF HEARTS
167	THE FIVE OF HEARTS
168	THE FIVE OF HEARTS
169	THE FIVE OF HEARTS
170	THE FIVE OF HEARTS
171	THE FIVE OF HEARTS
172	THE FIVE OF HEARTS
173	THE FIVE OF HEARTS
174	THE FIVE OF HEARTS
175	THE FIVE OF HEARTS
176	THE FIVE OF HEARTS
177	THE FIVE OF HEARTS
178	THE FIVE OF HEARTS
179	THE FIVE OF HEARTS
180	THE FIVE OF HEARTS
181	THE FIVE OF HEARTS
182	THE FIVE OF HEARTS
183	THE FIVE OF HEARTS
184	THE FIVE OF HEARTS
185	THE FIVE OF HEARTS
186	THE FIVE OF HEARTS
187	THE FIVE OF HEARTS
188	THE FIVE OF HEARTS
189	THE FIVE OF HEARTS
190	THE FIVE OF HEARTS
191	THE FIVE OF HEARTS
192	THE FIVE OF HEARTS
193	THE FIVE OF HEARTS
194	THE FIVE OF HEARTS
195	THE FIVE OF HEARTS
196	THE FIVE OF HEARTS
197	THE FIVE OF HEARTS
198	THE FIVE OF HEARTS
199	THE FIVE OF HEARTS
200	THE FIVE OF HEARTS

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818

— — — — —

[illegible]

Dh
450 CCCCCCCTCCCCCAGCCACCTGCGCCCTTCCCACTGCCG 519

161 IvsI,euAraG]nArI,euI,euAraSpA)aaSpaSpI,euG]nIvsArI,euA)IaVaI,Tyr 180

Db 520 AAGCTGCCTAAGCGGCTCCTCCGCGATGCCGATGACCTGCAGAACGCCCTGGCAGTGTAC 579

QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200

Db 580 CAGGCCGGGCCGAGGGCGCCCTCAGGCCATCCGCGAGCGCTGGG 639

QY 201 PROLeuValIGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220

Db 640 CCCCTGGTGGACAGGCCGCGTGGCGGCCCTGCACCTGTGGGCTCCCTGGCCGGCAG-CCG 698

QY 221 LeuGlnGluArgAlaGlnAla--TrpGlyGluArgLeuArgAlaArgMetGluGluMetG 240

Db 699 CTACAGAGCGGCCAGGCCCTTGGGGCGAGCGGATGCCGCCGGATGGAGGAGATGG 758

QY 240 IYSerArgThrArgAspArgLeuAspGluValLysGluGluInValAlaGluValArgAla 259

Db 759 GCCACCGGACCG--GAACCGCTGGACGAGGTGAAGGAGCAGGTGGCGGAGGTGGCGCC 815

RESULT 10
BI600563
LOCUS 603244936F1 NIH_MGC_96 811 bp mRNA linear EST 07-SEP-2001
DEFINITION Homo sapiens CDNA clone IMAGE:5287329 5',
mRNA sequence.
ACCESSION BI600563
VERSION BI600563.1 GI:15493502
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 811)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLM11726 row: g column: 10
High quality sequence stop: 783.
Location/Qualifiers
1..811
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5287329"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 148 a 245 c 316 g 102 t
ORIGIN
Alignment Scores:
Pred. No.: 2,7e-114 Length: 811
Score: 1185.00 Matches: 240
Percent Similarity: 98.37% Conservative: 2
Best Local Similarity: 97.56% Mismatches: 3
Query Match: 74.76% Indels: 1
DB: 13 Gaps: 0
US-09-827-854-19 (1-317) x BI600563 (1-811)
QY 1 MetlysaValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 74 ATGAAGGTTCTGTGGCGCTGCTGTCTGTCACATTCCTGGCAGATGCCAGCCAAAGGTG 133
QY 21 GtuglnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 134 GAGCAAGCGGTGGAGACAGACCGAGCGCGCCAGCCAGACCGAATGGGAGAGC 193
QY 41 GtlylnArGTrpGluLeuAlaLeuGlyArGpHeTrpAspTrpLysArgTrpValGlnThr 60
Db 194 GGCAGCGCTCGGAACGTGCACTGGGTGCTTTGGGATTAACGTGGCTGGGTGCAGACA 253
QY 61 LeuserGtuglnValGlnGluLeuLeuSerGlnValThrGlnGluLeuArgAla 80
Db 254 CTGTCTGAGCAGGTGACGAGGAGGAGCTGCTCCAGGTCCACCGAGGAAGTGAAGGCG 313

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu 100
Db 314 CTGATGAGCAGACCATATGAAGAGTGAAGGCTCAATATGGAACCTGGAGAACACTG 373
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 374 ACCCGGTGGCGGAGAACCGCGGACGCGTCTCCAAAGAGAGTCCAGCGCGCGAGCC 433
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyVal 140
Db 434 CGGTGGCGGCGACATGAGAGACTGTGGCGCGCTGTGTCACTACCGCGCGAGGTG 493
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValAlaGleuAlaSerHisLeuArg 160
Db 494 CAGCGCATGCTCGGCACAGACGACGAGAGCTGGCGGTCGCTGCCCTCCACCTGGCG 553
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 554 AAGCTGCTAAGCGGCTCTCCGCGATGCCATGCAGAGAGCGCTGGCACTGAC 613
QY 181 GlnAlaGlyAlaArgGluGlnGluValArgGlyLeuSerAlaIleArgGluArgLeu 200
Db 614 CAGCGCGGCGCGCGAGGCGCGAGCGCGCTCAGCGCATCCGAGCGCGCTGGGG 673
QY 200 YProLeuValGluGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyLys 220
Db 674 CCCCTGTGGAGACAGGCGCGCGTCCGCGCGCCACCTGTGGCTCCCTGGCGAGCACC 733
QY 220 oLeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetG1 240
Db 734 GCTACAGAGAGCGGCGCGCGCTGGCGGAGCGCTGCGCGCGAGATGAGAGCAATGGG 793
QY 240 ySerArgThrArgasp 245
Db 794 CAGCGGGAACCGCGA 809
RESULT 11
BG829472 845 bp mRNA linear EST 22-MAY-2001
LOCUS BG829472
DEFINITION Homo sapiens CDNA clone IMAGE:489112 5',
mRNA sequence.
ACCESSION BG829472
VERSION BG829472.1 GI:14177059
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 845)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLM1790 row: g column: 17
High quality sequence stop: 829.
Location/Qualifiers
1..845
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:489112"
/clone_lib="NIH_MGC_42"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pORF7; Site_1: XhoI;


```

QY 61 LeuSerGluGlnValGlnGluGluLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 253 CTGCTGAGAGAGGTGAGAGGAGGAGCTGCTCAGCTCCAGGTCACCCAGAACTGAGGGCG 312
QY 81 LeuMetAspLthrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluLeu 100
DB 313 CTGATGACAGACCATGAAAGGAGTTGAAAGCCCTACAAACCGAACTGGAGGAACAAC 372
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 373 ACCCGGGTGGCGAGAGAGACCGGCGCACTGCTCCAGAGAGCTGAGAGGGCGAGGCC 432
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyAlaVal 140
DB 433 CGGCTGGCGCGGACATGAGAGACGTGTGGCGCGCTGTGTGACATGACCGCGGAGGTG 492
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 493 CAGCGCAATGCTCGCGCAGAGCACCGAGAGCTCGGGTGGCTCGCTCCACTGCGC 552
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 553 AAGCTCGTGAAGGCTCTCCCGCGATGCCGATGACCTCGAAGAGCGCTGGCAGGTGAC 612
QY 181 GlnAlaGlyAlaArgGluGlnGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 613 CAGCGCGGGGCGCGGAGGGCGCGGCGGCTCAGCGCCATCCGCGCAGCGCTGGGG 672
QY 201 ProLeuValGluGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
DB 673 CCCCTGGTGGAAACAGGGCGCGTGGCGCGCCACTGTGGCTCGCTCGCGCAGCGCG 732
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyArgLeuArgAlaArgMetGluGluMet 239
DB 733 CTACAGAGAGCGGCGCCAGGCTGGGGCGAGCGGTG-CGCGCGCGGATGAGAGAGATG 788
RESULT 13
BI551475 919 bp mRNA linear EST 05-SEP-2001
LOCUS 603194314F1 NIH_MGC_95 Homo sapiens CDNA clone IMAGE:5274003 5',
DEFINITION mRNA sequence.
ACCESSION BI551475
VERSION BI551475.1 GI:15438787
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-remail.nih.gov
Tissue procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroaki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM11691 row: 1 column: 04
High quality sequence stop: 812.
FEATURES
Source
location/Qualifiers
1..919
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5274003"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"

```

```

/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
size-selected for average insert size 2.5 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI, National
Institutes of Health). Note: this is a NIH-MGC Library."
BASE COUNT 172 a 270 c 363 g 114 t
ORIGIN
Alignment Scores:
Pred. No.: 5,17e-110 Length: 919
Score: 1145.00 Matches: 246
Percent Similarity: 93.638 Conservative: 4
Best Local Similarity: 92.138 Mismatches: 16
Query Match: 72.24# Indels: 4
DB: 13 Gaps: 0
US-09-827-854-19 (1-317) x BI551475 (1-919)
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 73 ATGAAGCTTCTGTGGGCTCGTGTGTCACATTCCTGGCAGAGATGCCAGCGCAAGGTG 132
QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGluThrGluTrpGlnSer 40
DB 133 GAGCAAGCGGTGGAGACAGAGCCGAGACCCGACTCGCCAGCAGACCGAGTGGCAGAGC 192
QY 41 GlyGlnArgTrpLysLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTyrValGlnThr 60
DB 193 GGCACAGCGTGGGAATGCGACACGCGGTCTTTGGGATTACCTGGCGTGGGCGAGACA 252
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 253 CTGCTGACAGAGGTGAGAGAGGAGCTCAGCTCCAGCTCCAGCAAGCAAGTGAAGGGCG 312
QY 81 LeuMetAspLthrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGluLeu 100
DB 313 CTGATGACAGAGACCATGAAAGAGTTGAAGCCCTACAAATCGGAAGTGAAGAACTG 372
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 373 ACCCGGTGGCGAGAGAGACCGCGGCGCTCTCCAGAGAGCTGACAGCGCGCGAGGCC 432
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyAlaVal 140
DB 433 CGGCTGGCGCGGACATGAGAGACGTGTGGCGCGCTGTGTGACATGACCGCGGAGGTG 492
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 493 CAGGCCATGCTGGCGCAGAGCACCGAGAGCTGGGGTGGCTCGCTCCACTGCGC 552
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 553 AAGCTCGTGAAGGCTCTCCCGCGATGCCGATGACCTCGAAGAGCGCTGGCAGGTGAC 612
QY 181 GlnAlaGlyAlaArgGluGlnGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 613 CAGCGCGGGGCGCGCATGAGAGCGCGCTGCTCAGCGCCATCCGCGAGCGCTGGGG 672
QY 201 ProLeuValGluGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
DB 673 CCCCTGGTGGAAACAGGGCGCGGCGGCGCTGCTGCTCGCTCGCGCGGCGAGCCCT 732
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyArgLeuArgAlaArgMetGluGluMetGlu 240
DB 733 A-CAGGAGCGGGGCGCAGGCTGGGGCGAGCGGTGCGCGCGCGATGAGGAATAATGGG 790
QY 240 ySerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 791 CAGCGGAGCCCGGAGAGC-CTGGAGCAAGTGAAGACACGCTGGCGGAGAGG-TGGCCCAA 847

```


QY 260 sleuglgluglglalagl 266
|||||:::||||
Db 848 GCTGAGACGACGACGAG 866

RESULT 14
BM042228

LOCUS
DEFINITION BM042228 757 bp mRNA linear EST 07-NDV-2001
603616186F1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:557004 5',
mRNA sequence.
BM042228
VERSION BM042228.1 GI:16771495
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE 1 (bases 1 to 757)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cyabs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2021 row: k column: 21
High quality sequence stop: 757.

FEATURES
source
1. 757
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:557004"
/tissue_lib="NIH_MGC_112"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene)
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 134 a 229 c 296 g 98 t

ORIGIN

Alignment Scores:
Pred. No.: 4.39e-109 Length: 757
Score: 1135.00 Matches: 234
Percent Similarity: 99.16% Conservative: 1
Best Local Similarity: 98.73% Mismatches: 1
Query Match: 71.61% Indels: 2
DB: 13 Gaps: 0

US-09-827-854-19 (1-317) x BM042228 (1-757)

QY 1 MetlvalleutrrpalaalaaleuValThrPheleualaglycysglnalalysVal 20
|||||:::|||||
Db 50 ATGAAGTTCTGTGGCTGGCTGTGTCATCTCTGCGAGATGCCAGGCAAGTG 109

QY 21 GluglAlaValglutrrgluprogluarpgluaruglncglntrrglutrrgluSer 40
|||||:::|||||
Db 110 GAGCAACCGGTGGAGAGCGGAGCCGAGCTGCGCAGACGAGCGAGTGCAAGC 169

QY 41 G1yGlnArtrrpglualeualaleuG1yArpPheTrpAspTyrleuArgrTrpValGlnThr 60
|||||:::|||||
Db 170 GGCCAGCGCTGGAGACTGGAGTGGCTTTGGATTACCGCGCTGGTGACAGACA 229

QY 61 LeuSerGlucInValGlnGluglulLeuLeuSerSerGlnValThrGlnGlueAlaAla 80
|||||:::|||||
Db 230 CTGTCTAGCAGAGGTGCGAGAGAGCTCTCAGCTCCAGCTCACCCAGAACTGAGGGC 289

QY 81 LeuMetAspGluThrMetLysGluLeuValATyrLysSerGluLeuGluglulLeu 100
|||||:::|||||
Db 290 CTGATGACAGACCATGAAGAGCTTAAGGCTTACAAATCGGAACGTGGAGAACTG 349

QY 101 ThrProValAlaGluglulThrArGalaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
|||||:::|||||
Db 350 ACCCGGTGGCGAGAGACGGCGACGGCTGTCCAAAGAGCTGCAGCGCGCGAGGCC 409

QY 121 ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTrrArgGlyGluVal 140
|||||:::|||||
Db 410 CGCGTGGCGCGGACAGAGAGAGCTGTGGCGCGCTGTGGAGTACCCGCGAGGTG 469

QY 141 GluAlaMetLeuGlyGlnSerThrGluGluAlaValArgLeuAlaSerHisLeuArg 160
|||||:::|||||
Db 470 CAGGCAATGCTCGGCCAGACACCGAGAGACTGCGGGTGGCTCTCCACCTGCGC 529

QY 161 LysLeuArgGlnArGleuLeuArGAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||:::|||||
Db 530 AAGCTGGTAAGGGGCTCTCCGGATGCCGATGACCTGCAGAAAGCCCTGGCAGGTGATC 589

QY 181 GluAlaGluAlaArGluGluGluAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||:::|||||
Db 590 CAGGCGGGCGCGCGAGCGCGCGAGCGCGCTCAGCCCATCCGCGAGCCCTGGGG 649

QY 201 ProLeuValGluGlnGlyArgVal-ArgAlaIleThrValGlySerLeuAlaGlyGlnP 220
|||||:::|||||
Db 650 CCCCTGTGGACAGAGCGCCGCTGGCGCGCGCACCTGTGGCTCTCTGGCGCCAGCC 709

QY 220 oleuglgluarglglalaglAlaTrpGluArgLeuArgAlaArgMet 236
|||||:::|||||
Db 710 GCTACAGAGCGGCG-CAGGCTGGGGCGAGCGGCTCGCGCGGATG 757

RESULT 15
BG706129 907 bp mRNA linear EST 07-MAY-2001
BG706129 602669093F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4792030 5',
mRNA sequence.

LOCUS
DEFINITION BG706129.1 GI:13981169

ACCESSION BG706129

VERSION BG706129.1

KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE 1 (bases 1 to 907)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cyabs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10669 row: i column: 23
High quality sequence stop: 832.

FEATURES
source
1. 907
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4792030"
/tissue_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: Brain; Vector: pBluescriptR (modified

bluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 171 a 260 c 358 g 117 t 1 others

ALIGNMENT SCORES:

Pred. No.:	5,69e-108	Length:	907
Score:	1125.50	Matches:	243
Percent Similarity:	94.62%	Conservative:	3
Best Local Similarity:	93.46%	Mismatches:	10
Query Match:	71.01%	Indels:	6
DB:	12	Gaps:	2

US-09-827-854-19 (1-317) x BG706129 (1-907)

```
OY      1 MetLysValLeuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB      75 ATGAGGTTCTGTGGGCTGCTGCTGTCACATTCTGTCAGAGATGCCAGCAAGTG 134
OY      21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGlnPheGlnSer 40
DB      135 GAGCAAGCGGTGGAGACAGCGGAGCCCGAGCTGCCACAGACCCAGATGGCAGAGC 194
OY      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB      195 GGCACGCGCTGGGAACCTGCTGCTTTGGATTACCTGGCTGGCTGGCGAGACA 254
OY      61 LeuSerGlnGluValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB      255 CTGTCTGACAGGTGACGAGGAGCTGCTCAGCTCCAGTCCAGCAGCACTGAGGGCG 314
OY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
DB      315 CTGATGGAGACACCATGAGAGAGTTGAAGGCTTACAAATCCGAACTGGAGACACTG 374
OY      101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGlu-LeuGlnAlaAlaGlnAl 120
DB      375 ACCCGGTGGCGAGAGACGCGGCGACCGCTGTCCAAAGAGCTGCAGCGCGCGAGGC 434
OY      120 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB      435 CCGGCTGGGCGGACATGAGAGCGTGTGCGCGCTGTGTGCACTACCGCGCGAGGT 494
OY      140 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB      495 GCAGGCGCATGCTCGGCGACAGACACCGAGAGAGTGGGGTGGCTCGCTCCACCTGGC 554
OY      160 GlySerLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB      555 CAAGCTGGGTAAAGCGGCTCTCCGCGATGCCATGACCTGCAGAAAGCGCTGGCAGTGA 614
OY      180 rGlnAlaGlyAlaArgGluGluAlaGluArgGlyLeuSerAlaIleArgGluArgLeu-G 200
DB      615 CCAGGCGCGGGGCGCGGAGGCGCGGAGGCGGCTCAAGCGCCATCCGAGCGGCTGGG 674
OY      200 LysProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnP 220
DB      675 GCCCTTGGTGAACAGGCGCGCGTGGCGCGCCACTGTGGGCTCCCTGCGNCGCGCACG 734
OY      220 rLeuGlnGlnAlaGlnAlaTrpGlyGlyArgLeuArgAlaArgMetGlnGlnMetG 240
DB      735 CGCTACAGAGCGCGCGAGCTG---GGCGAGCGGTGGCGCGCGGATGAGAGATGG 791
OY      240 LysArgTrpThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArg 258
DB      792 GCAG-CGGACCGG---ACGCTGAGAGAGTG-AAAGAGCAGTGGGGAAGTGTGCGC 842
```

Search completed: March 14, 2003, 20:14:18
Job time : 1293.69 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:08:17 ; Search time 178.728 Seconds

(without alignments)
3994.237 Million cell updates/sec

Title: US-09-827-854-15

Perfile score: 1585
Sequence: 1 MKVLMALLVFLAGCQAKV.....VEKYQAAVTSAPSPSDNH 317

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp
-O=/cgn2.1/USPTO/US09827854/runat_11032003_101609_27476/app.query.fasta.1.3576
-DB=N.Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNIT=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi
-LIST=45 -DOCALLGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptco -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854.GCGN.1.1.1201 &runat_11032003_101609_27476 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY NEGSCORES=0 -MATT LONGLOG -DEV_TIMCOUT=120
-WARN_TIMCOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :

N.Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Prod. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1585	100.0	954	24	AAD26035	Human apolipoprote
2	1585	100.0	1110	7	AAN60409	Human apolipoprote
3	1585	100.0	1147	22	ABA83113	Apolipoprotein E o
4	1585	100.0	1156	22	AAF84315	Human ApoE3 coding
5	1585	100.0	1156	24	AAD22048	Human apolipoprote
6	1581	99.7	1156	24	AAD22052	Human apolipoprote
7	1577	99.5	1156	22	AAF84314	Human ApoE2 coding
8	1577	99.5	1156	24	AAD22049	Human apolipoprote
9	1577	99.5	1156	24	AAD22051	Human apolipoprote
10	1576	99.4	1157	17	AAT06957	Human apolipoprote
11	1576	99.4	1157	24	ABN95746	Human apolipoprote
12	1576	99.4	1157	24	ABK64514	Gene #2244 used to
13	1576	99.4	1157	24	ABL65450	Human benign prost
14	1575	99.4	1110	6	AAN50450	Lung cancer relate
15	1573	99.2	1156	22	AAF84316	Sequence encoding
16	1573	99.2	1156	24	AAD22047	Human ApoE4 coding
17	1570	99.1	1156	24	AAD22050	Human apolipoprote
18	1481.5	93.5	1279	22	AA522437	Human apolipoprote
19	1462.5	92.3	1107	19	AA575756	Human cDNA encodin
20	1396.5	88.1	9360	24	ABL31915	Human ApoE genomic
21	1396.5	88.1	10716	24	AAD26034	Human apolipoprote
22	1391	87.8	3805	20	AA209526	Human Apo E genom
23	1391	87.8	3805	20	AAD209526	Human Apo E genom
24	1369.5	86.4	10716	24	AAD26108	Human Apo E genom
25	1163.5	73.4	965	24	AAD32081	Human alpha-1-anti
26	1163.5	73.4	5617	24	AAD32077	Human albumin prom
27	1163.5	73.4	6026	24	AAD32075	Nucleotide sequenc
28	1138	71.8	1126	19	AAV29159	ApoB4x2 protease
29	985	62.1	936	15	AAQ69101	Human ApoE4x2 CDN
30	985	62.1	936	17	AAT18070	Partial human apol
31	984	62.1	660	18	AAT69792	Human ApoE4 CDNA.
32	947	59.7	597	17	AAT18068	ApoE4L protease cd
33	938	59.2	597	15	AAO69099	Human cDNA encodin
34	884	55.8	1381	22	AA522673	Lung cancer associ
35	867.5	54.7	786	21	AAF18114	EST clone CP147.
36	678	42.8	600	20	AAV89595	Gene #2492 used to
37	651	41.1	407	24	ABK34238	Colon adenocarcino
38	625	39.4	478	24	ABN95994	Thyroid cancer rel
39	625	39.4	478	24	ABL62679	Human diagnostic a
40	625	39.4	478	24	ABL67340	Human apolipoprote
41	594	37.5	499	22	AA503049	Human secreted pro
42	514	32.4	330	12	AAQ11980	Human EST-derived
43	504	31.8	405	21	AAQ02139	Human apolipoprote
44	485	30.6	345	22	AAH98479	Human apolipoprote
45	419	26.4	260	21	AAA40342	Human apolipoprote

ALIGNMENTS

RESULT 1

AAD26035
AAD26035 standard; CDNA; 954 BP.

26-MAR-2002 (first entry)

Human apolipoprotein E (APOE) CDNA.

Human; antilipemic; neuroprotective; nootropic; genetic variant; APOE;
apolipoprotein E; haplotype; familial dysbetalipoproteinemia; therapy;
genotyping; type III hyperlipoproteinemia; Alzheimer's disease; SNP;
atherosclerosis; single nucleotide polymorphism; chromosome 19q13.2; ss.

Homo sapiens.
Location/Qualifiers

FT	CDS	1..954	
FT		/*tag= a	
FT	variation	/product= "Human APOE protein"	
FT		replace (13, C)	
FT		/*tag= a	
FT	variation	/standard_name=	"Single nucleotide polymorphism (SNP) "
FT		replace (31, G)	
FT		/*tag= b	
FT	variation	/standard_name=	"Single nucleotide polymorphism (SNP) "
FT		replace (364, A)	
FT		/*tag= c	
FT	variation	/standard_name=	"Single nucleotide polymorphism (SNP) "
FT		replace (388, C)	
FT		/*tag= d	
FT	variation	/standard_name=	"Single nucleotide polymorphism (SNP) "
FT		replace (498, G)	
FT		/*tag= e	
FT	variation	/standard_name=	"Single nucleotide polymorphism (SNP) "
FT		replace (526, T)	
FT		/*tag= f	
FT	variation	/standard_name=	"Single nucleotide polymorphism (SNP) "
FT		replace (622, A)	
FT		/*tag= g	
FT		/standard_name=	"Single nucleotide polymorphism (SNP) "
XX	MO200179234-A2.		
XX	25-OCT-2001.		
XX	16-APR-2001; 2001MO-US12303.		
XX	14-APR-2000; 2000US-197188P.		
XX	(GENA-) GENAISSANCE PHARM INC.		
XX	Choi JY, Kilem SE, Koshy B, Lee HH;		
XX	WPI: 2002-075064/10.		
XX	P-PSDB; AAE15158.		
XX	Genotyping human apolipoprotein gene of individual for determining		
XX	haplotype of individual, involves determining identity of nucleotide		
XX	pair at specific polymorphic sites for two copies of gene -		
XX	Claim 26; Fig 2; 78pp; English.		
XX	The patent discloses novel genetic variants of human apolipoprotein		
XX	E (APOE) gene. The invention also relates to compositions and methods		
XX	for haplotyping and/or genotyping the APOE gene. The haplotyping		
XX	methods of the invention are useful for improving the efficacy and		
XX	reliability of several steps in the discovery and development of		
XX	drugs for treating diseases associated with APOE activity, e.g.		
XX	familial dysbetalipoproteinaemia, type III hyperlipoproteinaemia,		
XX	atherosclerosis, and Alzheimer's disease. They are useful to validate		
XX	APOE as a candidate agent for treating a specific condition or disease		
XX	predicted to be associated with APOE activity and in the design of		
XX	clinical trials of candidate drugs for treating a specific condition		
XX	or disease predicted to be associated with APOE activity. Genotyping		
XX	or haplotyping methods are useful to screen for compounds targeting		
XX	APOE to treat a specific condition or disease associated with APOE		
XX	activity. The present sequence is a cDNA encoding human APOE protein.		
XX	APOE gene is located on chromosome 19q13.2.		
XX	Sequence 954 BP; 168 A; 290 C; 377 G; 119 T; 0 other;		
XX	Alignment Scores:		
XX	Pred. No.:	1..02e-114	Length: 954
XX	Score:	1585.00	Matches: 317
XX	Percent Similarity:	100.00%	Conservative: 0
XX	Best Local Similarity:	100.00%	Mismatches: 0
XX	Query Match:	100.00%	Indels: 0
XX	DB:	24	Gaps: 0

US-09-827-854-15 (1-317) x AAD26035 (1-954)	
QY 1 MetLysValLeuTPtAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20	
Db 1 ATGAAGGTTCTGTGGCTGCGCTGTGTCTGCATTTCCGTCGAGATGCCAGGCCAAGTGG 60	
QY 21 GlnGlnAlaValaGlnThrGlnProGlnProGlnLeuLeuArgGlnGlnGlnThrGlnTrpGlnSer 40	
Db 61 GAGCAAGCGGCTGGAGACAGAGACCGGAGCCCAAGCTGGCCGACAGACCGAAGTGGCAAGC 120	
QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlnYArgPheTrpAspTyrLeuArgTrpValaGlnThr 60	
Db 121 GGCCAGCGCTGGAGACTGGCACTGGGTGCTTTGGCATTTACCTGCGCTGGGTGCACACA 180	
QY 61 LeuSerGlnGlnValaGlnGlnGlnLeuLeuLeuSerGlnValaThrGlnGlnLeuArgAla 80	
Db 181 CTGTGTGAGCAGAGTGCAGAGAGAGCTGTCTCAGCTCCAGGTCCACAGAACTAGAGGCG 240	
QY 81 LeuMetAspGlnThrMetLysGlnLeuLeuYsaIatYrLysSerGlnLeuGlnGlnLeu 100	
Db 241 CTGATGGACAGAGACCATAGAGAGTTGAAGGCTTCAAAATGGAACTGGAGGAACAACCTG 300	
QY 101 ThrProValaAlaGlnGlnThrArgAlaLeuSerLysGlnLeuGlnAlaGlnAla 120	
Db 301 ACCCGGCTGGGGAGAGAACCGGGCAAGCGCTGTCCAAGAGAGCTGACGCGGCGCCAGCGCC 360	
QY 121 ArgLeuGlnYAlaAspMetGlnLysPvalaCysGlnYArgLeuValaGlnTrpArgGlnVala 140	
Db 361 CGGCTGGGCGGAGATGAGAGAGCTGTGGCGGCTGTGGTCAATACCGCGGAGAGTG 420	
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuLeuArgValaArgLeuAlaSerHisLeuArg 160	
Db 421 CAGGCCATGTCTGGCCAGAGACCGAGAGAGCTGGGGGTGCGCTCCGCTCCACCTCGCC 480	
QY 161 LysLeuArgLysArgGlnLeuLeuArgAspAlaAspLysPheGlnLysArgLeuAlaValaTyr 180	
Db 481 AAGCGCTTAAGCGGCTCTCCGCGATGCCATACCTGCAGAAAGCCCTGGCAGTCTAC 540	
QY 181 GlnAlaGlnYAlaArgGlnGlnYAlaGlnYArgGlnYLeuSerAlaIleArgGlnYArgLeuGlnY 200	
Db 541 CAGCGCGGGGCGCGGAGAGGCGCGAGGCGGCGCTCAGAGCCCATCGCGAGCGGCTGGGG 600	
QY 201 ProLeuValaGlnGlnGlnYArgValaArgAlaAlaThrValaGlnSerLeuAlaGlnPro 220	
Db 601 CCCCGTGGTGAACAGAGCGCGGCTCGGGCCGCCACTGTGGCTCCCTGGCGGCGCACCG 660	
QY 221 LeuGlnGlnYArgAlaGlnAlaIleTrpGlnYArgLeuArgAlaArgMetGlnGlnMetGlnY 240	
Db 661 CTACAGAGAGCGGCGCCAGGCTGGGGCGAGCGGCTGCCCGCGCGGATGGAGAGATGGCC 720	
QY 241 SerArgThrArgAspArgLeuAspGlnValaLysGlnGlnValaAlaGlnValaArgAlaLys 260	
Db 721 AGCCGAGACCCCGACCGCGCTGCAGAGGTGAAGAGCAGGCGGCGGAGTGGCGGCCAAG 780	
QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgGlnGlnAlaGlnAlaPheGlnAlaArgLeuLys 280	
Db 781 CTGGGAGGAGCAGGCCCAAGATATAGCTTCGAGGCGCAGAGCGCTTCCAGGCGCGCTTAAG 840	
QY 281 SerTrpPheGlnProLeuValaGlnAspMetGlnArgGlnTrpAlaGlnYLeuValaGlnLys 300	
Db 841 AGCTGCTTCGAGCCCTGTGTGTGAAGACATGACGCGCCAGTGGGCGGGGTGTGGAGAAAG 900	
QY 301 ValGlnAlaAlaValaGlnYThrSerAlaAlaProValProSerAspAsnHis 317	
Db 901 GTGCGAGCTGCCGTGGGCGACAGCGCGCGCCCTGTGCCACAGCGCAATCTAC 951	
RESULT 2	
AAN60409	
ID AAN60409 standard; DNA, 1110 BP.	
XX AAN60409;	
DT 01-JAN-1980 (first entry)	

XX New apoe humanized mammalian cell useful for screening for agents
 PT useful for treating or preventing Alzheimer's disease and
 PT arteriosclerosis -
 XX
 XX Disclosure: Page 13-14; 22pp; Japanese.
 XX
 CC The present invention relates to an Apoe humanised mammalian cell. The
 CC present sequence is the coding sequence for human ApoE3, which was used
 CC in the method of the present invention. The ApoE humanised mammalian cell
 CC can be used for screening for agents useful for treating or preventing
 CC Alzheimer's disease and arteriosclerosis.
 XX
 SQ Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other;

Alignment Scores:
 Pred. No.: 1,266-114 Length: 1156
 Score: 1585.00 Matches: 317
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-827-854-15 (1-317) x AAF84315 (1-1156)

OY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 61 ATGAAGGTTCTGTGGGCTGCTGGCTGGCTGACATTCCTGCGAGATGCCAGGCCAAGGTG 120
 OY 21 GlnGlnAlaValAlaGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 121 GAGCAAGCGGTGAGAGACAGCGCGGAGCCGAGCTGCGCCAGACAGACCGAGTGGCAGAGC 180
 OY 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB 181 GGCCAGCGGTGGAGTGCACCTGGCTGCTTTGGAGTACTGCGCGGTGGCAGACA 240
 OY 61 LeuSerGlnGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB 241 CTGTCTGAGCAGGTGAGAGAGAGCTGCTCAGCTCCAGGTACCCAGGAGTGAAGGCG 300
 OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluLeu 100
 DB 301 CTGATGAGACAGACCATGAGAGATTGAAGGCTTCAATCGAATCGAGAGACAGACACTG 360
 OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 361 ACCCGGTGGCGAGAGAGCGGGGACGCGTGTCCAAGAGACTGCAGCGCGCGCAGGCGC 420
 OY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 DB 421 CGGCTGGCGCGGACATGAGAGAGCTGTGCGCGCTGTGCTCAGTACCGCGCGCAGGGTG 480
 OY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 481 CAGGCCATGCTGGCGCAGAGCACCGAGAGTCTGGGTCGCTGCCCTCCACCTCGCGC 540
 OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 DB 541 AAGCTGCGTAAAGCGGCTCTCGCATGTGACCTGCAGAAACGCGCTGCGACGTAC 600
 OY 181 GlnAlaGlyAlaArgGlyGlyValArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 601 CAGGCGGGGCGCGGAGGCGCGCAGCGCGCTCAGCGCCATCCGAGCGCGCTGGGG 660
 OY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyLysPro 220
 DB 661 CCCCTGTGTGAACAGAGCGCGCTGGCGCGCCACTGTGGGTCCCTGGCGCGCAGCGC 720
 OY 221 LeuGlnGluArgAlaGlnAlaThrPglGlyGluArgLeuArgAlaArgMetGluMetGly 240
 DB 721 CTACAGGAGCGGCGCCAGGCGCTGGGGCGAGCGGCTGCGCGCGGATGAGAGATGGGC 780

OY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 DB 781 ACCCGGACCGCGGACCGCTGTGACAGTGAAGAGACAGTGGCGAGGTGGCGCGCAAG 840
 OY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 DB 841 CTGAGAGAGACAGCCAGCAGATACGCTTCAGCGCGAGGCTTCAGAGCCGCGCTCAAG 900
 OY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 DB 901 ACCTGTTTCAGACCCCTGTGTAAGACATGACAGCCGCACTGGCGCGGCTGTGTGAAG 960
 OY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 961 GTGCAGGCTGCGGTGGGACACAGCGCGCGCTGTGTGCCAGGACATTCAC 1011

RESULT 5
 AAD22048
 ID AAD22048 standard; DNA; 1156 BP.
 XX
 XX AAD22048;
 AC
 XX 12-FEB-2002 (first entry)
 DT
 XX Human apolipoprotein E (apoE) isoprotein, apoE3 DNA.
 DE
 XX Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;
 KW hypertriglyceridaemia; low density lipoprotein; LDL; ds.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 61..1014
 FT /*tag= a
 FT /product= "Human apoE isoprotein, apoE3"
 FT sig_peptide 61..114
 FT /*tag= b
 FT mat_peptide 115..1011
 FT /*tag= c
 FT /product= "Mature human apoE isoprotein, apoE3"
 FT
 PN MO200177136-A1.
 PD 18-OCT-2001.
 XX
 XX 06-APR-2001; 2001WO-US11358.
 PF
 XX 06-APR-2000; 2000US-0544386.
 PR 04-OCT-2000; 2000US-0679088.
 PR 05-APR-2001; 2001US-0827854.
 XX
 XX (KOSP-) KOS PHARM INC.
 PA (UYBO-) UNIV BOSTON.
 XX
 XX Zannis VI, Kypros KE;
 PI WPI: 2002-010885/01.
 DR P-PSDB: AAE13294.
 DR
 XX New apolipoprotein E polypeptide and nucleic acid, useful for lowering
 PT cholesterol, delaying the onset of or treating atherosclerosis in
 PT mammal, without inducing hypertriglyceridaemia -
 XX
 PS Claim 14; Page 81; 91pp; English.
 XX
 XX The present sequence is a human apolipoprotein E (apoE)
 CC isoprotein, apoE3 DNA. The apoE lipoproteins are useful for
 CC lowering cholesterol, delaying the onset of atherosclerosis,
 CC treating or regressing atherosclerosis without inducing
 CC hypertriglyceridaemia, in a mammal lacking an endogenous,
 CC normally functioning apoE gene or low density lipoprotein (LDL)
 CC receptor or is at risk for developing atherosclerosis due to
 CC accumulation of lipoprotein remnants in the bloodstream or having


```
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
    |||
Db 421 CGGCTGGGCGGACATGAGGAGACGTGTGGCGCGCTGTGACGTACCGCGGAGAGTG 480
QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
    |||
Db 481 CAGGCCATGCTGGCGCAGACCGAGAGAGCTGGGGTGGCTCGCTCCGCTCCACCTGCGC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
    |||
Db 541 AAGCTGTGTAAGGGGCTCCCGCGATGCCGATGACCTGACAGAAGTCCCTGGCAGTGTAC 600
QY 181 GlnAlaGlyAlaArgGlnGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
    |||
Db 601 CAGGCCGGGCGCGCAGCGCGCGGACCGCGGCTCAGCCCATCCCGCAGACCGCTGGGG 660
QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
    |||
Db 661 CCGCTGTGTGACAGAGGGCGCGGTGGCGGCCACACTGTGGGCTCTCGCGCGCGCACCG 720
QY 221 LeuGlnGluArgAlaGlnAlaTyrGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
    |||
Db 721 CTACAGAGAGCGGGCGCCAGGCTGGGGGAGCGGCTGGCGCGCGGATGAGAGATGGCG 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
    |||
Db 781 AGCGCGGACCGCGACCGCTGTGACGAGGTGACAGAGAGCGAGGTGGCGGAGCTGGCG 840
QY 261 LeuGlnGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
    |||
Db 841 CTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
    |||
Db 901 AGCTGTGTGAGAGCGCTCGGTGGAGAACATGACAGCGCAGTGGGGCGGGCTGTGGAGAA 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
    |||
Db 961 GTCCAGAGCTCCGTGGGACACAGCGCGCGCTGTGCCAGCAGCAATCTAC 1011
RESULT 8
ID AAD22049 standard; DNA; 1156 BP.
AC AAD22049;
XX
DT 12-FEB-2002 (first entry)
XX
DE Human apolipoprotein E (apoE) isoprotein, apoE2 DNA.
XX
KM Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;
XX hypertriglyceridaemia; low density lipoprotein; LDL; ds.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 61..1014
FT sig_peptide 61..114
FT mat_peptide 115..1011
FT /tag= a
FT /product= "Human apoE isoprotein, apoE2"
FT /tag= b
FT /tag= c
FT /product= "Mature human apoE isoprotein, apoE2"
XX
PN WO200177136-A1.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001MO-US11358.
XX
PR 06-APR-2000; 2000US-0544386.
```

```
PR 04-OCT-2000; 2000US-0679088.
PR 05-APR-2001; 2001US-0827854.
XX
PA (KOSP-) KOS PHARM INC.
XX (UYBO-) UNIV BOSTON.
PI Zannis VI, Kypros KE;
XX WPI; 2002-010885/01.
DR P-PSDB; AAE13295.
XX
PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT cholesterol, delaying the onset of or treating atherosclerosis in
PT mammal, without inducing hypertriglyceridaemia.
XX
PS Claim 14; Page 81-82; 91pp; English.
XX
CC The present sequence is a human apolipoprotein E (apoE)
CC isoprotein, apoE2 DNA. The apoE lipoproteins are useful for
CC lowering cholesterol, delaying the onset of atherosclerosis,
CC treating or regressing atherosclerosis without inducing
CC hypertriglyceridaemia, in a mammal lacking an endogenous,
CC normally functioning apoE gene or low density lipoprotein (LDL)
CC receptor or is at risk for developing atherosclerosis due to
CC accumulation of lipoprotein remnants in the bloodstream or having
CC a defect in remnant removal.
XX
SQ Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;
XX
Alignment Scores:
Pred. No.: 5,3e-114 Length: 1156
Score: 1577.00 Matches: 316
Percent Similarity: 99.68% Conservative: 0
Best Local Similarity: 99.68% Mismatches: 1
Query Match: 99.50% Indels: 0
DB: 24 Gaps: 0
US-09-827-854-15 (1-317) x AAD22049 (1-1156)
QY 1 MetLysValLeuTrpAlaAlaLeuValAlaThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||
Db 61 ATCAAGGTTCTGTGGGCTGCGTGTGTCACATTCTCGCAGAGATGCCAGGCCAAGGTG 120
QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnTrpGlnSer 40
    |||
Db 121 GAGCAAGCGGTGGAGAGACAGAGCGCGAGCCGAGCTCGCGCAGACAGACCGAGTGGCAGAC 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
    |||
Db 181 GGCACAGCGCTGGGACACTGACAGCGGCTCTTGGGATTACTCGCGCTGGGTGACAGACA 240
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
    |||
Db 241 CTGTCTGAGCAGAGTGCAGAGAGAGAGCTGTGCTACGTCTCCAGCTACCCAGAACAGAGCGCG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnGlnLeu 100
    |||
Db 301 CTGATGACAGACCATGAAGAGGTGAAGGCTTACAAATCGGAACCTGAGAGAACACTG 360
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGluLeuGlnGlnAlaGlnAla 120
    |||
Db 361 ACCCGGTGGGAGAGAGAGCGCGGCGGCTCTCCAAAGAGCTGACGCGCGCGCGCGCG 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
    |||
Db 421 CGGCTGGGCGGACATGAGGAGACGTGTGGCGCGCTGTGACGTACCGCGGAGAGTG 480
QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
    |||
Db 481 CAGGCCATGCTGGCGCAGACCGAGAGAGCTGGGGTGGCTCGCTCCGCTCCACCTGCGC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
    |||
```


OY 301 ValGlnAlaIaValGlyThrSerAlaAlaProSerAspAsnHis 317
|||||
DB 962 GTCAGAGGCTGCGGTGGGACACAGCCGCCCTGTGCTCCAGGACATTCAC 1012

RESULT 11

ABN95746
ID ABN95746 standard; DNA; 1157 BP.

AC ABN95746;

DT 13-AUG-2002 (first entry)

DE Gene #2244 used to diagnose liver cancer.

Gene: liver cancer; ds: hepatocellular carcinoma; hepatotropic;
metastatic liver tumour; cytostatic; expression profile; disease state;
disease progression; drug toxicity; drug efficacy; drug metabolism.

OS Homo sapiens.

PN W0200229103-A2.

PD 11-APR-2002.

PF 02-OCT-2001; 2001WO-US30589.

PR 02-OCT-2000; 2000US-237054P.

PA (GENE-) GENE LOGIC INC.

PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

DR WPI; 2002-426119/45.

PT Diagnosing and detecting the progression of liver cancer,
hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -

PS Claim 1; SEQ ID NO 2244; 298bp; English.

CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumor in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytosolic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcr_sequences.

XX Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;

Alignment Scores:

Pred. No.: 6,34e-114 Length: 1157
Score: 1576.00 Matches: 315
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.43% Indels: 0
DB: 24 Gaps: 0

US-09-827-854-15 (1-317) x ABN95746 (1-1157)

OY 1 MetLysValLeuTrpAlaIaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
DB 62 ATGAAGGTTCTGTGGGCTGCTGTGCATTCCTGCGAGATGCCAGGCCAAGGTG 121

OY 21 GlnAlaIaValaIaGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
|||||
DB 122 GAGCAAGCGGTGGACACAGAGCCGAGCGCTGCGCCAGACACAGCCAGTGGCAGAC 181

OY 41 GlyIlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
|||||

DB 182 GGCACAGCGCTGGAACTGGACACTGGTGGCTTTTGGATTAACCTCGCTGGTGGACACA 241

OY 61 LeuSerGluGlnValaIaGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||

DB 242 CTGTCTGAGCAGTGCAGAGAGAGCTGCTCAGCTCCAGTCCACCAAGAACCTGAGGGCG 301

OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGluGlnLeu 100
|||||

DB 302 CTGATGAGACGAGACCATGATGAGAGCTTGAAGCCTTCAAAATCCGAATCTGGAGGACACTG 361

OY 101 ThrProValaIaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaGlnAla 120
|||||

DB 362 ACCCGGTAGCGAGAGAGAGCGGGGACGGCTGTCCAAAGAGACTGCAGACGGCGCAGGCC 421

OY 121 ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnTrpArgGlyVala 140
|||||

DB 422 CGGCTGGGCGGACATGAGAGAGCTGTGGCGCCCTGTGTCACTACCTGCGGCGAGGTG 481

OY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValaArgLeuAlaSerHisLeuArg 160
|||||

DB 482 CAGGCCATGCTGCGCCAGACAGCACCGAGAGCTGCGGGTGTGCTGCTCCACCTGCGC 541

OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTrp 180
|||||

DB 542 AAGCTGCGTAAAGCGGCTCTCCGCGATCCCGATGACCTGCAGAACCGCTGCGACGTAC 601

OY 181 GlnAlaGlyAlaArgGluGluValaGluArgGlyLeuSerAlaIlaArgGluArgLeuGly 200
|||||

DB 602 CAGCGCGGGGCGCGGAGGCGCCGAGCGGCTTCACGCGCATCCGAGCGGCTGTGGGG 661

OY 201 ProLeuValaIaGluGlnGlyArgValaArgAlaIaThrValaGlySerLeuAlaGlyGlnPro 220
|||||

DB 662 CCCCTGTGTGAACAGAGCGCGCTGCGGCGGCGGCTGTGTGCTGCTGCGGCGGCGGCA 721

OY 221 LeuGlnGluArgAlaGlnAlaIaTrpGlyGluArgLeuArgAlaArgMetGlnGluMetGly 240
|||||

DB 722 CTACAGAGAGCGGCGCCAGCGCTGCGGCGAGCGGTGCGGCGGATGGAGAGATGGGC 781

OY 241 SerArgTrpArgAspArgLeuAspGluValaLysGluGlnValaIaGluValaArgAlaLys 260
|||||

DB 782 AGTCGAGCCCGGACCGCTGTGACAGGTGAAGGAGCAGGTGCGGAGCTGCGCGCCCAAG 841

OY 261 LeuGlnGluGlnAlaGlnIlnLeuArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
|||||

DB 842 CTGGAGGAGCAGGCCAGCAGATACGCTGCGAGCGCGGAGGCTTCCAGGCGCGCTCAAG 901

OY 281 SerTrpPheGluProLeuValaIaAspMetGlnArgGlnTrpAlaGlyLeuValaGluLys 300
|||||

DB 902 AGCTGTTGAGAGCCCTGCTGTGAAGACATGACAGCCCAAGTGGCGGCTGTGTGAAGAAG 961

OY 301 ValGlnAlaIaValaIaGlyThrSerAlaAlaProValaProSerAspAsnHis 317
|||||

DB 962 GTGAGGCTGCTGTGGGACAGCAGCGCCGCCCTGTGCTCCAGGACATTCAC 1012

RESULT 12

ABK64514
ID ABK64514 standard; DNA; 1157 BP.

AC ABK64514;

DT 18-JUN-2002 (first entry)

DE Human benign prostatic hyperplasia gene #409.

Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.

OS Homo sapiens.

XX WO200212440-A2.
PN
XX
PD 14-FEB-2002.
XX
PF 07-AUG-2001; 2001WO-US24708.
XX
PR 07-AUG-2000; 2000US-223323P.
PR 05-JUN-2001; 2001US-0873319.
XX
PA (GENE-) GENE LOGIC INC.
PA (NIBS) JAPAN TOBACCO INC.
XX
PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
PI WPI: 2002-257476/30.
XX
XX
PT Identifying drugs for and diagnosing benign prostatic hyperplasia, by
PT detecting expression levels of one or more genes in prostate cells from
PT patient that are differentially regulated compared to normal prostate
PT cells -
XX
PS Disclosure: Page 239-240: 444pp: English.
XX
CC The invention relates to a method of diagnosing (I) the onset or
CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
CC or identifying an agent that modulates the onset or progression of BPH.
CC The method is based on changes in gene expression in BPH tissue isolated
CC from patients exhibiting different clinical states of prostate
CC hyperplasia as compared to normal prostate tissue. (I) comprises
CC detecting the expression levels of one or more genes in prostate cells
CC from the subject that are differentially regulated compared to normal
CC prostate cells. (II) comprises preparing a first gene expression profile
CC of BPH cells or BPH-like cell population, exposing the cells to the
CC agent, preparing a second gene expression profile of the agent exposed
CC cells, and comparing the first and second gene expression profiles.
CC (I) is useful for diagnosing the onset or progression of BPH. (II) is
CC useful for identifying an agent that modulates the onset or progression
CC of BPH. The methods are useful to present information identifying
CC the expression level in a tissue or cells, by comparing the expression
CC level of genes given in the specification in the tissue or cells to the
CC level of expression of gene in the database, and displaying the
CC expression levels of at least one gene in the tissue or cell sample
CC compared to the expression level in BPH. Agents using (II) are useful for
CC treating BPH or prostate cancer. ABK64106-ABK64860 represent human
CC benign prostatic hyperplasia gene sequences of the invention.
XX
SQ Sequence 1157 BP: 212 A; 370 C; 426 G; 149 T; 0 other:

Alignment Scores:
Pred. No.: 6.34e-114 Length: 1157
Score: 1576.00 Matches: 315
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.43% Indels: 0
DB: 24 Gaps: 0

US-09-827-854-15 (1-317) x ABK64514 (1-1157)
QY 1 MetIysValLeuTrrpAlaLeuLeuValThrPheLeuAlaGlycysGlnAlaLysVal 20
DB 62 ATGAAGGTTCTGTGGCTCGTGTGTCACATTCTCGCAGGAGCCAGGCCAAGG 121
QY 21 GluGlnAlaValaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 122 GAGCAAGCGGTGGAGACAGACGCCGAGCCGAGCTGCGCAGCAGACCGAGTGGCAGGC 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 182 GGCAGAGCGCTGGGAACTGGCAGTGGCTTTGGGATTACTGCGCGGGGCGAGACA 241
QY 61 LeuSerGluGlnValaGlnGluGluLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||

DB 242 CTGCTGAGCAGGTGCGAGGAGAGACTGCTACGCTCCCAAGTCACCCAGAACTGAGCGCG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
DB 302 CTGATGGACGACACCATGAAAGAGTTGAAGGCTTACAAATCCGAACTGGAGAACACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 362 ACCCGGTTAGCGGAGAGACGCGGGCAGCGCTGTCCAAAGAGCTGCAGACGCGCGAGCGCC 421
QY 121 ArgLeuGluAlaAspMetLysPvalLysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 422 CGGCTGGCGCGGACATGGAGACGTTGTGGCGCGCTGTGCAGTACCGCGAGGTG 481
QY 141 GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 482 CAGCCATGCTCGGCCAGACACCGAGAGCTCGGGTCCGCTCCGCTCCACCTGCGCC 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 542 AAGCTCGTAAAGCGCTCTCCGCGATCCCGATGACCTGCAGAAAGCGCTGGCAGTGTAC 601
QY 181 GlnAlaGlyAlaArgGluGluGluAlaGluArgGlyLeuSerAlaThrLeuArgLysLeuGly 200
DB 602 CAGCGCGGGCGCCGCGAGGCGCGCGCTCAGCGCCATCCGCSAGCGCTGGGG 661
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
DB 662 CCCCTGTGTGAAACAGGGCGCGCTGGCGCGCCACTGTGGCTCCTGGCGCGCCAGCGG 721
QY 221 LeuGlnGluArgAlaGlnAlaAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
DB 722 CTACAGACGCGGGCCAGCGCTGGGGCGAGCGGCTCGCGCGCGATGAGAGATGGGCG 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 782 AGTCAGACCCGCGACCGCTGTGACGAGGTGAAGAGACAGTGGCGAGGTGGCGCCAGG 841
QY 261 LeuGluGlnGlnAlaGlnGlnThrArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 842 CTGGAGAGACAGGCCAGCAGATACGCTCGAGCGGAGCGGCTTCAGGCCCGCCCTCAAG 901
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 902 AGCTGTTGTGAGCCCTGTTGTGAAGACATGACAGCCGCACTGGGCGCGTGTGGAGAG 961
QY 301 ValGlnAlaAlaValaGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 962 GTGCAAGCTGCGGTGGGACACAGCCGCCCTGTGCTCCAGCGACATAC 1012

RESULT 13
ABL65450
ID ABL65450 standard; DNA: 1157 BP.
XX
AC ABL65450;
XX
XX
DE 15-MAY-2002 (first entry)
XX
XX Lung cancer related gene sequence SEQ ID NO:3787.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
XX gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX

PR	05-JUN-2000;	2000US-209473P.
PR	05-JUN-2000;	2000US-209531P.
PR	18-SEP-2000;	2000US-233133P.
PR	18-SEP-2000;	2000US-233617P.
PR	20-SEP-2000;	2000US-234009P.
PR	20-SEP-2000;	2000US-234034P.
PR	20-SEP-2000;	2000US-234052P.
PR	22-SEP-2000;	2000US-234509P.
PR	22-SEP-2000;	2000US-234567P.
PR	25-SEP-2000;	2000US-234923P.
PR	25-SEP-2000;	2000US-234924P.
PR	25-SEP-2000;	2000US-235077P.
PR	25-SEP-2000;	2000US-235082P.
PR	25-SEP-2000;	2000US-235134P.
PR	25-SEP-2000;	2000US-235280P.
PR	26-SEP-2000;	2000US-235637P.
PR	26-SEP-2000;	2000US-235638P.
PR	27-SEP-2000;	2000US-235711P.
PR	27-SEP-2000;	2000US-235720P.
PR	27-SEP-2000;	2000US-235840P.
PR	27-SEP-2000;	2000US-235863P.
PR	28-SEP-2000;	2000US-236028P.
PR	28-SEP-2000;	2000US-236032P.
PR	28-SEP-2000;	2000US-236033P.
PR	28-SEP-2000;	2000US-236034P.
PR	28-SEP-2000;	2000US-236109P.
PR	28-SEP-2000;	2000US-236111P.
PR	29-SEP-2000;	2000US-236842P.
PR	29-SEP-2000;	2000US-236891P.
PR	02-OCT-2000;	2000US-237172P.
PR	02-OCT-2000;	2000US-237173P.
PR	02-OCT-2000;	2000US-237278P.
PR	02-OCT-2000;	2000US-237294P.
PR	02-OCT-2000;	2000US-237295P.
PR	02-OCT-2000;	2000US-237315P.
PR	03-OCT-2000;	2000US-237342P.
PR	03-OCT-2000;	2000US-237598P.
PR	03-OCT-2000;	2000US-237604P.
PR	03-OCT-2000;	2000US-237606P.
PR	03-OCT-2000;	2000US-237608P.
PR	01-NOV-2000;	2000US-244867P.
PR	01-NOV-2000;	2000US-245084P.
PA	(AVALON PHARM.	
PI	Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;	
PI	Soppet DR, Weaver Z;	
XX	WPI: 2002-188264/24.	
DR	XX	
XX	XX	
PT	Screening for anti-neoplastic agent involves exposing cells to a	
PT	chemical agent to be tested for anti-neoplastic activity, and	
PT	determining a change in expression of a gene of a signature gene set	
XX	Claim 1; SEQ ID 3787; 44pp; English.	
PS	The present invention describes a method (M1) for screening for an	
XX	anti-neoplastic agent. The method involves exposing cells to a chemical	
CC	agent to be tested for anti-neoplastic activity, determining a change in	
CC	expression of at least one gene (I) of a signature gene set, where (I)	
CC	comprises a sequence (S) selected from 8447 sequences (given in ABU61664	
CC	to ABU70110), or is at least 95% identical to (S), where a change in	
CC	expression is indicative of anti-neoplastic activity. (I) has cytostatic	
CC	activity and can be used in gene therapy. M1 can be used for screening	
CC	an anti-neoplastic agent, and can be used for producing a product which	
CC	is the data collected with respect to the anti-neoplastic agent as a	
CC	result of M1, and the data is sufficient to convey the chemical	
CC	structure and/or properties of the agent. M1 can be used in the	
CC	treatment of cancer such as colon, breast, stomach, lung, thyroid,	
CC	oesophagal, ovarian, kidney, prostate or pancreatic cancer,	
CC	adenocarcinoma, carcinoma, great cell cancer, infiltrating ductal cancer	
CC	infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine	
CC	carcinoma, papillary carcinoma and Wilms' tumour.	

XX	Sequence	1157 Bp; 212 A; 370 C; 426 G; 149 T; 0 other;
SQ		
Alignment Scores:		
Pred. No.:	6,34e-114	Length: 1157
Score:	1576.00	Matches: 315
Percent Similarity:	99.37%	Conservative: 0
Best Local Similarity:	99.37%	Mismatches: 2
Query Match:	99.43%	Indels: 0
DB:	24	Gaps: 0
US-09-827-854-15 (1-317) x ABL65450 (1-1157)		
QY	1 MettlyValluettPpAlaAlaLeuValThrPhleuAlaGlyCySglnAlaIysVal	20
Db	62 ATGAAGGTTCTCTGGGCTGCGCTTGTGTGCATTTCCCTGGCAGGATGCCAGGCCAAGGTG	121
QY	21 GluGlnAlaValAlaGluThrGluProGluProGluLeuAArgGlnGlnThrGluTrpGlnSer	40
Db	122 GAGCAAGCGGCTGGACACAGAGCCGAGCCCAAGCTGCCGCCAGCAGACGAGTGGCAGAGC	181
QY	41 GlyGlnAArgTrpGluLeuAlaLeuGlyAArgPheTrpAspTrpLeuAArgTrpValGlnThr	60
Db	182 GGGCAGCGCGTGGGAAGCTGGCACTGGGTGCTTTTGGATTAACGTGGCGGTGGGTGACAGCA	241
QY	61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuAArgAla	80
Db	242 CTGTCTGAGCAGGTGTCAGAGAGAGGTGCTCATGCTCCCAAGTCACCCCAAGACTGAGGGCG	301
QY	81 LeuMeTAPGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGluGlnLeu	100
Db	302 CTGATGAGCAGAGACCATTAAGAGATTGAAGGCTTACAAATGTGGAAGTGGAGAAACAATG	361
QY	101 ThrProValAlaGluGluGluThrArgAlaAArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	362 ACCCGGTAGCGGAGAGACAGCGGGGACAGCGGTGTCTCAAGAGAGTGCAGACGGCGCACGGCC	421
QY	121 ArgLeuGlyAlaAspMetGluAspValCysGlyAArgLeuValGlnTrpAArgGlyGluVal	140
Db	422 CGGCTGGGCGCGGACATGAGAGAGGTGTGCGCGCGCGCTGAGTACCCGCGGAGAGTG	181
QY	141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuAArgValAArgLeuAlaSerHisLeuAArg	160
Db	442 CAGGCGCATGCTCGGCGAGAGACCGAGAGAGCTGGGGGTGCGCTCTCCACCTGCGC	541
QY	161 LysLeuAArgLysAArgLeuLeuAArgAspAlaAspAspLeuGlnLysAArgLeuAlaValTyr	180
Db	542 AAGCTGCTAAGCGGCTCTCCCGGATCCGATACCTGCAGAGAGCCCTGGCAGTGTAC	601
QY	181 GlnAlaGlyAlaAArgGluGlyAlaGlnAArgGlyLeuSerAlaIleAArgGluAArgLeuGly	200
Db	602 CAGCGCGGGGCGCGGAGAGGCGCGGAGCGCGGCTCAAGCGCATTCGCGAGCGGCTGGG	661
QY	201 ProLeuValGluGlnGlyAArgValAArgAlaIleAThrAValGlySerLeuAlaGlyGlnPro	220
Db	662 CCCCGTGGTGAACAGAGGCGCGGTGGGGCGCCCACTGGGGCTCCCTGAGCGGCGCACCG	721
QY	221 LeuGlnGluAArgAlaGlnAlaIleTrpGlyGlyAArgLeuAArgAlaAArgMetGluMetGly	240
Db	722 CTACAGGAGCGGCGCCACAGCGCTGGGGCGAGCGGTGCGCGCGCGCATGAGAGATGCGC	781
QY	241 SerAArgTrpAArgAspAArgLeuAspGlyValLysGluGlnValAlaGlnValAArgAlaLys	260
Db	842 CTGGAGGAGCAGGCGCCACCAATATACGCTTCAGAGCCGAGAGGCTTCCAGGCGCCGCTCAAG	901
QY	261 LeuGluGluGlnAlaGlnGlnIleAArgLeuGlnAlaGlnAlaIleAArgLeuLys	280
Db	901 AGCTGGAGACCCGCGACCGCTTGACAGAGTGAAGAGCAAGGTGAGGCGGAGTGGTGAAG	961

QY 301 ValGlnAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 962 GTGCAGAGCGCGTGGGACACAGCGCCGCTGTGCCAGGACAAATCACC 1012

RESULT 14
AAN50450
ID AAN50450 standard; DNA; 1110 BP.
XX
AC AAN50450;
XX
DT 09-JAN-1992 (first entry)
XX
DE Sequence encoding human apolipoprotein E.
XX
KM Hyperlipaemia; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 15..968
FT /tag= a
FT mat_peptide 69..965
FT /tag= b

XX JP60118189-A.
XX
XX PD 25-JUN-1985.
XX
XX PF 29-NOV-1983; 83JP-0224980.
XX
XX PR 29-NOV-1983; 83JP-0224980.
XX
XX PA (MITU) MITSUBISHI CHEM IND KK.
XX
XX DR WPI: 1985-188003/31.
XX P-PDB; AAPS1204.
XX
XX PT DNA fragment - contg. DNA which codes human apolipoprotein E for
XX treatment of hyperlipaemia.
XX
XX PS Claim 3; Page 484; 8pp; Japanese.
XX
XX CC The sequence may be used to produce the apolipoprotein E, useful in the
XX treatment of hyperlipaemia.
XX
SQ Sequence 1110 BP; 198 A; 355 C; 414 G; 143 T; 0 other;

Alignment Scores:
Pred. No.: 7,24e-114 Length: 1110
Score: 1575.00 Matches: 315
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.37% Indels: 0
DB: 6 Gaps: 0

US-09-827-854-15 (1-317) x AAN50450 (1-1110)

QY 1 MetLysValLeuThrAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaTyrVal 20
DB 15 ATGAAAGTTCTGTGGGCTGCTGCTGTCACATTCTTCCGACAGATGCCAGGCAAGTTC 74

QY 21 GluGlnAlaValAlaGlnThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 75 GAGCAAGCGGTGGAGACAGAGCCGAGAGCTGCGCAGACAGACCGAGTGGCAGAGC 134

QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 135 GGCACACCGCTGGAACTGCACCTGGGCTGCTTTGGATTACCTGCGTGGGTGCAGACA 194

QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 195 CTGTCTGACACAGCTGACGAGGAGCTGTCACTCCAGGTACCCAGGAAGTGAAGGCG 254

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
DB 255 CTGATGAGACAGACCATGAAGAGTTGAAGGCTTACAAATCGGAACGTGAGCAACATCG 314

QY 101 ThrProValAlaGlnGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
DB 315 ACCCGGTGGGAGGAGACGCGGCGACGCTGTCCAAAGAGCTGAGCGCGCCAGCGCC 374

QY 121 ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 375 CGGCTGGCGCGGACATGAGAGACGTGTGCGCGCGCTGTGTGACAGTACCGCGGAGGTTC 434

QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 435 CAGGCCATCTCGCGCAGACACCGAGAGCTGCGGCTGCCCTCCCTCCACCTGCCGC 494

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 495 AAGCTCGTAAAGCGGCTCTCCGCGATGCCGATGACCTCAGAAAGCGCTGGCAGTGTAC 554

QY 181 GlnAlaGlyAlaArgGlnGluGlyAlaGlyArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 555 CAGCGCGGGGCGCGAGGGCGCGAGCGCGGCTCAGCGCCATCCGAGCGCGCTGGGG 614

QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 615 CCCCTGTGTGAACAGGCGCGCGTGTGCGCGCGCCACTGTGTGGCTCCCTGCCCGCGAGCCG 674

QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
DB 675 CTACAGAGAGCGGCGCCAGGCTGTGGGCGAGCGGCTCCGCGCGGTGAGAGATGGGC 734

QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 735 AGCGGAGACCGCGAGCGCGCTGAGCAGGTGAAGAGACAGTGGCGAGCGCGCGCAAG 794

QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 795 CTGGAGAGACAGGCGCCAGCAGATACCCCTGACAGCGCGGCTTCCAGGCGCGCTCAAG 854

QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 855 AGCTGTTTGAAGCCCTGTGTGAAGACATGCACGCCCACTGGGCGCGCTGTGGGAAG 914

QY 301 ValGlnAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 915 GTGCAGAGCGCGTGGGACACAGCGCCGCTGTGCCAGGACAAATCACC 965

RESULT 15
AAF84316
ID AAF84316 standard; CDNA; 1156 BP.
XX
AC AAF84316;
XX
XX DT 21-JUN-2001 (first entry)
XX
XX DE Human Apoe4 coding sequence.
XX
XX DE Human; Apoe4; Alzheimer's disease; arteriosclerosis; ss.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
FT CDS 61..1014
FT /tag= a
FT /product= "Human Apoe4"

XX JP2001017028-A.
XX
XX PD 23-JAN-2001.
XX
XX PF 28-APR-2000; 2000JP-0128919.
XX

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:17:52 ; Search time 34.7343 Seconds
(without alignments)
2798.866 Million cell updates/sec

Title: US-09-827-854-15
Perfect score: 1585
Sequence: 1 MKVLMALVTLFLGCAKAV.....VEKVQAAVGTSAAPVPSDNH 317

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n_model -DEV=xlp
-Q=/cgn2_1/USPFO.spool/US09827854/rnatac_11032003_101611_27510/app_query.fasta_1.3576
-DB=issued.patents_NA -QPM=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR.SCORE=100 -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptlo -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09827854.cgn2_1.1.274.@rnatac_11032003_101611_27510 -NCPU=6 -ICPU=3
-NO.XLPRY -NO.MMAP -LARGEOURR -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-MARK.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCRTUS.COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1576	99.4	1157	1	US-07-709-949-1
2	1138	71.8	1126	4	US-08-949-155-5
3	1138	71.8	1126	4	US-09-819-964-5
4	1002	63.2	4267	4	US-08-949-155-51
5	1002	63.2	4267	4	US-09-819-964-51
6	984	62.1	660	2	US-07-726-306A-28
7	511	32.2	330	1	US-07-849-389-6
8	374	23.6	252	3	US-08-617-256-24
9	374	23.6	252	4	US-09-287-141-24
10	374	23.6	252	4	US-09-431-613-24
11	374	23.6	252	4	US-09-504-245-24
12	374	23.6	252	4	US-09-287-682-24

13	374	23.6	252	4	US-09-287-679-24	Sequence 24, Appl
14	374	23.6	252	4	US-09-397-766-24	Sequence 24, Appl
15	374	23.6	252	4	US-09-287-681-24	Sequence 24, Appl
16	374	23.6	252	4	US-09-495-444-24	Sequence 24, Appl
17	177.5	11.2	842	1	US-08-952-796-1	Sequence 1, Appl
18	174.5	11.0	801	1	US-07-959-946-4	Sequence 4, Appl
19	174.5	11.0	801	1	US-08-333-577-4	Sequence 4, Appl
20	174.5	11.0	801	5	PCT-US92-08634-4	Sequence 4, Appl
21	159	10.0	5661	4	US-08-938-105-2	Sequence 2, Appl
22	157.5	9.9	964	4	US-08-448-606-5	Sequence 5, Appl
23	154.5	9.7	863	1	US-08-448-606-7	Sequence 7, Appl
24	154	9.7	3256	2	US-08-968-751-3	Sequence 3, Appl
25	153	9.7	1879	4	US-09-750-580-2	Sequence 2, Appl
26	147.5	9.3	603	4	US-08-952-796-14	Sequence 14, Appl
27	146	9.2	8789	1	US-08-328-254-5	Sequence 5, Appl
28	146	9.2	10136	1	US-08-353-700-2	Sequence 2, Appl
29	146	9.2	10136	5	PCT-US95-16216-2	Sequence 2, Appl
30	145	9.1	4868	1	US-08-139-937-12	Sequence 12, Appl
31	145	9.1	4868	5	PCT-US93-11310-12	Sequence 12, Appl
32	145	9.1	6306	1	US-08-195-487-3	Sequence 3, Appl
33	145	9.1	6306	5	PCT-US93-06160-3	Sequence 3, Appl
34	144	9.1	1771	2	US-08-533-669A-7	Sequence 7, Appl
35	144	9.1	1771	2	US-08-511-872-1	Sequence 1, Appl
36	144	9.1	1771	4	US-09-183-861-7	Sequence 7, Appl
37	144	9.1	1771	4	US-09-022-765-7	Sequence 7, Appl
38	144	9.1	13121	4	US-08-961-527-126	Sequence 126, App
39	143.5	9.1	11236	4	US-07-853-913-1	Sequence 1, Appl
40	141	8.9	516	4	US-09-183-861-34	Sequence 34, Appl
41	141	8.9	516	4	US-09-022-765-34	Sequence 34, Appl
42	141	8.9	6306	1	US-08-466-390-3	Sequence 3, Appl
43	141	8.9	6306	1	US-08-470-950-3	Sequence 3, Appl
44	141	8.9	6306	1	US-08-467-781-3	Sequence 3, Appl
45	141	8.9	6306	2	US-08-483-924-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-07-709-949-1
Sequence 1, Application US/07709949
Patent No. 5472858
GENERAL INFORMATION:
APPLICANT: Attie, Alan D
APPLICANT: Gretch, Daniel G
APPLICANT: Sturley, Stephen L
APPLICANT: Beckage, Nancy E
TITLE OF INVENTION: Production of Recombinant Proteins in
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: P. O. Box 2113
CITY: Madison
STATE: WI
COUNTRY: USA
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/709,949
FILING DATE: 19910604
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 9629691801
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 1:

Db 8/9 CTCCTCAAGGCTGGTTCGAGCCCTCTGGTGGAGACATACGGCCCACTGGGCCGGGCTG 93

111

111


```

QY 38 TrrglnSerGlycInAtrgrrglnLeuAlaLeuGlyAtrpPhetrrpAspyrrleuAtrgrrp 57
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 159 TGGCAGGAGGAGCCAGCCCTGGAGAGCAGCCCTGGAGGCTTCTGGATTACCGCCCTGG 218
QY 58 ValGlnThrleuSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 77
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 219 GTCCAGTCCCTGTCTGACCAAGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 278
QY 78 LeuArgAlaLeuMetAspGlnThrMetGlyGlnLeuGlyAlaTrpGlySerGlnGlnGln 97
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 279 CTGACGAGAGCTGATPAGAGAGACATGAGAGAGTGAAGGCTACCGCGAGAGGAGGAG 338
QY 98 GlnGlnLeuThrProValAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 117
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 339 GCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 398
QY 118 AlAGlnAlaArgLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 399 GCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 458
QY 138 GlyGlnValGlnAlaMetLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 157
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 459 AGCAGAGTGCACAACATGTTGGCCAGACACCGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 518
QY 158 HisLeuArgGlySerLeuArgGlySerLeuArgGlySerLeuArgGlySerLeuArgGly 177
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 519 CACCTCGGCAAGGTCGCGCAAGCGGCTGCTCGGCGACACCGAGGAGGAGGAGGAGGAGGAG 578
QY 178 AlaValTrpGlnAlaGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 579 GCCGTGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 638
QY 198 ArgLeuGlyProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 217
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 639 CGCCTCGGCGCCCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 698
QY 218 GlyGlnProLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 237
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 699 GGCAGACCCCTGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 758
QY 238 GluMetGlySerArgTrpArgAspArgLeuAspGlnGlnGlnGlnGlnGlnGlnGlnGln 257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 759 GAGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 818
QY 258 ArgAlaLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 277
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 819 CGCACAAGAGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 878
QY 278 ArgLeuGlySerTrpPheGlnProLeuValGlnAspMetGlnArgGlnGlnGlnGlnGln 297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 879 CTCCTCAAGGCTGTGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 938
QY 298 ValGlnValGlnAlaValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 939 GTGGAAGAGTGAAGTGGGCTGTGAGCATTAAGTCTCTCCACCTCTGCGCCAGTGAAT 998

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/949,155
FILING DATE: Concurrently Herewith
CLASSIFICATION: 800
Prior APPLICATION DATA:
APPLICATION NUMBER: US 60/027,338
FILING DATE: 11-OCT-1996
Prior APPLICATION DATA:
APPLICATION NUMBER: US 60/046,094
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hieber, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TANK:177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 4267 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-949-155-51

Alignment Scores:
Pred. No.: 2,21e-87 Length: 4267
Score: 1002.00 Matches: 218
Percent Similarity: 58.43% Conservative: 42
Best Local Similarity: 48.99% Mismatches: 46
Query Match: 63.22% Indels: 140
DB: 4 Gaps: 5
US-09-827-854-15 (1-317) x US-08-949-155-51 (1-4267)
QY 1 MetGlyValLeuThrPheAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
    ||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2448 ATGCGTGTGAGTACGCTCGCT-----GCAGATGCGCGAGACAGAGGC 2489
QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArg-----GlnGlnThrGln 37
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2490 GAGCCGCGG-----CCGCCGCGAGAGTGCACGTGNGTGGAGGAGGAGCCCAAG 2537
QY 38 TrrglnSerGlyGlnAtrgrrglnLeuAlaLeuGlyAtrpPhetrrpAspyrrleuAtrgrrp 57
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2538 TGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2597
QY 58 ValGlnThrleuSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 77
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2598 GTCCAGTCCCTGTCTGACCAAGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2657
QY 78 Leu----- 78
Db 2658 CT-GACGTAAGTGGCCACCCGACTCCCGCGCGCGCGCGCGCGCGCGCGCTGA 2716
QY 78 ----- 78
Db 2717 CCCTCTGGCAACCGTGTGTTGTGACCCCTCAGGCTCCACCGCTCGGGTTCTCTTG 2776
QY 78 ----- 78
Db 2777 TCCTGTGCCAACCTCTTGGGGGTCTGGGTCTGTCTCTTTCTTTCTCTCTCTTT 2836
QY 78 ----- 78
Db 2837 TTGGGGGAAAAAAGCTTTTCTTTTCTTTTCTTTTCAATTGACTCATGCTTCTTTCC 2896
QY 78 ----- 78

```



```

Db 2837 TTGGGGGGAAAAAATTTTCTTTCTTTCTTTCTTCAATTGAGTCAATGATGCTTCTTCTTC 2893
Oy 78 ----- 78
Db 2897 ATCTTGACCTCTGCTGCTTCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2956
Oy 78 ----- 78
Db 2957 AATCTGTGGACAGCTCTGCGCATCGCCAGTCAAGAGCCCTCTTCTCCCTTACGCCC 3016
Oy 79 -----ArgAlaLeuMetAspGluThrMetLysGluIleuLysAlaThr 92
Db 3017 CCGGCCCTCTCTGCGCCAGGAGGATGATGAGAGAGATGAAGAGATGAAGAGAGTGAAGCCCTAC 3076
Oy 93 LysSerGluLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 112
Db 3077 CGCGAGGAGCTGGAGGGCGAGCTGGGGCCCTGATCCAGCAAGAACAGCAGGCGGCTGCTGCC 3136
Oy 113 LysGluLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 132
Db 3137 AAGGAGCTGCAGGGGGCGGCGAGCCCGCTGGGGCCGACATGAGAGACCTGGCAGACCC 3196
Oy 133 LeuValGlnThrArgGlyGluValGlnAlaMetLeuGlyGlnSerThrGluGluIleuValArg 152
Db 3197 TTGGGCTCTACCGCCAGCGAGGTGCACAAATGTTGGCGAGACACCGAGAGAGTGGG 3256
Oy 153 ValAlaGluLeuAlaSerThrLysLeuArgLysLeuArgLysLeuLeuValAspAlaAspAsp 172
Db 3257 AGCCGCTGGCTTCCACACCTGCGCAGCTGCGCACACGGCTGCTCCGGCAGCACCGAGAC 3316
Oy 173 LeuGluLysArgGluLeuAlaValThrGlnAlaGlyAlaArgGluGluValAlaGluValArg 192
Db 3317 CTGCAGGAAGCGCTGGCGCGGTACACAGCGCGGCTGCGCGAGCGCGCGCAGCGCAGCGCTG 3376
Oy 193 SerAlaIleArgGluArgLeuGlyGlyProLeuValGluGlnGlyArgValArgAlaAlaThr 212
Db 3377 AGCGCCCTCGCGAGCGCCCTCGGGCCCTGCTGTGAGACAGGGCCATTGCGCGCCACCC 3436
Oy 213 ValGlySerLeuAlaGlyGlnProLeuGluGlnAlaArgAlaIleThrGlyGluValArg 232
Db 3437 CTGAGTACACAGGCGCGCGCCAGCGCTGCGCGACCGCGCGGAAGCTGGGGCGACAGACTG 3496
Oy 233 ArgAlaIleArgMetGluMetGlySerArgThrArgAspArgLeuAspGluValLysIle 252
Db 3497 CGCGAGCGCGCTGGAGGAGATGGGACCGGACCGCGGACCGCTGTGATGATGCTGTAG 3556
Oy 253 GlnAlaIleArgValAlaArgAlaLysLeuGluGlnAlaGlnAlaIleArgLeuGlnAla 272
Db 3557 CAGCTGGAGGAGGTCGCCACCAAGTGGAGAGACAGACGAGCAGCACTGCGCTCGAGGCC 3616
Oy 273 GlnAlaIleArgGlnAlaArgLeuLysSerThrPheGluProLeuValGluAspMetGlnArg 292
Db 3617 GAGGAGTATGCCAGCGCTTCCCAAAAGCTGTGTGACCTCTGTGGAAGACATACATAGCGGC 3676
Oy 293 GlnThrPheAlaGlyLeuValGluLysValGlnAlaIleAlaValGly---ThrSerAlaIleArg 311
Db 3677 CAGTGGCGCGCGCTGGTGGAGAGATGACAGTGGCGCTGACATGAAGCTCTCCACACTCT 3736
Oy 312 ValProSerAspIle 316
Db 3737 GCGGCCAGCTGATTAAT 3751

RESULT 6
US-08-726-306A-28
: Sequence 28, Application US/08726306A
: Patent No. 5958684
: GENERAL INFORMATION:
: APPLICANT: van Leeuwen, Frederik Willem
: APPLICANT: Burbach, Johannes Peter Henri
: APPLICANT: Grosveld, Franklin G.
: TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
: NUMBER OF SEQUENCES: 189
: CORRESPONDENCE ADDRESS:

```

```

1 ADDRESS: Banner & Witcoff, Ltd.
2 STREET: 1 Financial Center
3 CITY: Boston
4 STATE: MA
5 COUNTRY: US
6 ZIP: 02111
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: Diskette, 3.50 Inch, 1.44 Mb storage
9 OPERATING SYSTEM: PC-DOS/MS-DOS
10 SOFTWARE: WordPerfect 6.1
11 CURRENT APPLICATION DATA:
12 APPLICATION NUMBER: US/08/726,306A
13 FILING DATE: 02-Oct-1996
14 PRIOR APPLICATION DATA:
15 APPLICATION NUMBER: GB 95/20080.4
16 FILING DATE: 02-Oct-1995
17 PRIOR APPLICATION DATA:
18 APPLICATION NUMBER: US 60/009,832
19 FILING DATE: 01-Jan-1996
20 ATTORNEY/AGENT INFORMATION:
21 NAME: Williams, Ph.D., Kathleen M.
22 REGISTRATION NUMBER: 34,380
23 REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
24 TELECOMMUNICATION INFORMATION:
25 TELEPHONE: (617) 345-9100
26 TELEFAX: (617) 345-9111
27 INFORMATION FOR SEQ ID NO: 28:
28 SEQUENCE CHARACTERISTICS:
29 LENGTH: 660 base pairs
30 TYPE: nucleic acid
31 STRANDEDNESS: double
32 TOPOLOGY: linear
33 MOLECULE TYPE: cDNA
34 US-08-726-306A-28
35
36 Alignment Scores:
37 Pred. No.: 9,84e-87 Length: 660
38 Score: 984.00 Matches: 197
39 Percent Similarity: 98.99% Conservative: 0
40 Best Local Similarity: 98.99% Mismatches: 2
41 Query Match: 62.08% Indels: 0
42 DB: 2 Gaps: 0
43
44 US-09-827-854-15 (1-317) x US-08-726-306A-28 (1-660)
45
46 QY 1 MettysValLeuTrpPalaIaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20
47 |||||||
48 Db 62 ATGAAGGTTCTGTGGCGCTTGCTGCTGACATTCTGCGAGATGCCAGGCCAAGTG 122
49
50 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
51 |||||||
52 Db 122 GAGCAAGCGGTGGAGACAGACC CGAGCCGAGCTGCGCCAGCAGACCGAGTGCAAGC 181
53
54 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspPyrIleuArgTrpValGlnThr 60
55 |||||||
56 Db 182 GGCAGCGCTGGGAACGGCAGCTGGTCTGTTGGGATTACCGCCCTGGGTGCAGACA 241
57
58 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
59 |||||||
60 Db 242 CTGTCTGTGACAGGGCGAGAGAGACTGCTCAGCTCCCAAGTACCACCAAGAACTAGGGCG 301
61
62 QY 81 LeuMetAspGluTrpMetIysGluLeuIysAlaTrpIysSerGluLeuGlnGluGlnLeu 100
63 |||||||
64 Db 302 CTGATGGACGACGACCATGGAAGATTTGAAGGCCCTCAAAATCGGAACCTGAGGAACAAC 361
65
66 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerIysGluLeuGlnAlaIaGlnAla 120
67 |||||||
68 Db 362 ACCCGCGTAGCGGAGAGACCGCGGACGGCTGCCAAGGAGCTGACAGACGGCGCAGGCC 421
69
70 QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
71 |||||||
72 Db 422 CGGCTGGGCGCGGACATGAGAGAGAGCTGTGGCGCGCTGTGTGCAATGACCGCGCGAGGTG 481

```


MOLECULE TYPE: CDNA
US-09-431-613-24

Alignment Scores:
Pred. No.: 2.63e-28 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.60% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15 (1-317) x US-09-431-613-24 (1-252)

QY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
DB 2 GCACGGCTGTCCAAAGAGCTGCAGCGCGCGCGCGCTGGCGGACATGGAGAGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGCGC-CGCTGTGTGACGTACCGCGCGCGCGCATGCTCGCGCAGAGACAC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuArg 168
DB 121 GAGGAGCTGCGGGTGGCTCGCTCCACCTGCGCAAGCTGCTTAAGCGCTCTCTCCGC 180

QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGTCCCTGGCAGTGTACCAAGCGCGGCGCCCGAGAGGCGCC 240

QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTC 252

RESULT 11
US-09-504-245-24
Sequence 24, Application US/09504245
Patent No. 6221605
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCaulliffe LLP
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/504.245
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002J
TELECOMMUNICATION INFORMATION:
TELEPHONE: 858-450-8400
TELEFAX: 858-587-5360
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-504-245-24

Alignment Scores:
Pred. No.: 2.63e-28 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.60% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15 (1-317) x US-09-504-245-24 (1-252)

QY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
DB 2 GCACGGCTGTCCAAAGAGCTGCAGCGCGCGCGCGCTGGCGGACATGGAGAGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGCGC-CGCTGTGTGACGTACCGCGCGCGCGCATGCTCGCGCAGAGACAC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuArg 168
DB 121 GAGGAGCTGCGGGTGGCTCGCTCCACCTGCGCAAGCTGCTTAAGCGCTCTCTCCGC 180

QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGTCCCTGGCAGTGTACCAAGCGCGGCGCCCGAGAGGCGCC 240

QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTC 252

RESULT 12
US-09-287-682-24
Sequence 24, Application US/09287682
Patent No. 6235478
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCaulliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002E
TELECOMMUNICATION INFORMATION:

APPLICATION NUMBER: US/09/397,766
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-20021
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-397-766-24

Alignment Scores:
Pred. No.: 2,63e-28 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.60% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15 (1-317) x US-09-397-766-24 (1-252)

OY 109 AlaArgLeuSerLySGluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCGAGCGCGCGCGCGCGCGCTGCGGACATGAGAGAC 61
OY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGCGC-CGCTGTGTGAGTACCGCGCGCGAGGTGCAAGCCATCTCGGCCAGAGCAC 120
OY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGCGGCTCGCTCCCTCCACCTGCGCAAGCTGCGTAAGCGGCTCTCCGC 180
OY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGluAla 188
DB 181 GATGCCGATGACCTGCAGAGTCTCCCTGGAGTGTACCAAGCGCGGCGCGCGAGGGCGCC 240
OY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCGCTC 252

RESULT 15
US-09-287-681-24
Sequence 24, Application US/09287681
Patent No. 6277573
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,681
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-287-681-24

Alignment Scores:
Pred. No.: 2,63e-28 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.60% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15 (1-317) x US-09-287-681-24 (1-252)

OY 109 AlaArgLeuSerLySGluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCGAGCGCGCGCGCGCGCTGCGGACATGAGAGAC 61
OY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGCGC-CGCTGTGTGAGTACCGCGCGCGAGGTGCAAGCCATCTCGGCCAGAGCAC 120
OY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGCGGCTCGCTCCCTCCACCTGCGCAAGCTGCGTAAGCGGCTCTCCGC 180
OY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGluAla 188
DB 181 GATGCCGATGACCTGCAGAGTCTCCCTGGAGTGTACCAAGCGCGGCGCGCGAGGGCGCC 240
OY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCGCTC 252

Search completed: March 14, 2003, 20:18:27
Job time : 39.7343 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 17:32:58 ; Search time 65.8753 Seconds
(without alignments)
3377.460 Million cell updates/sec

Title: US-09-827-854-15
Perfect score: 1585
Sequence: 1 MKVLMALVTLFLAGCAKV.....VEKVOAAVGTSAFVPSDNH 317

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 501302 segs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_p2n.model
-O=/cgn2_1/USPTO.spool/US09827854/runat_11032003_101612_27578/app_query.fasta.1.3576
-DB=Published.Applications_NA -OPMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100
-THR.MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09827854.ecgn_1_1.298.6runat_11032003_101612.27578
-NCPU=6 -ICPU=3 -NO_XL_PXY -NO_MAP -LARGEOUTERT -NEG.SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published.Applications.NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PTI_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCRTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES				Description	
Result No.	Score	Query Match	Length DB ID		
1	1585	100.0	1156 9	US-09-870-759-129	Sequence 129, App
2	1585	100.0	1156 9	US-09-802-640-17	Sequence 17, Appl
3	1585	100.0	1156 10	US-09-827-854-8	Sequence 8, Appl
4	1585	100.0	1291 12	US-10-044-090-454	Sequence 454, App

5	1581	99.7	1156 10	US-09-827-854-12	Sequence 12, Appl
6	1577	99.5	1156 10	US-09-827-854-9	Sequence 9, Appl
7	1577	99.5	1156 10	US-09-827-854-11	Sequence 11, Appl
8	1576	99.4	1157 10	US-09-954-456-760	Sequence 760, App
9	1576	99.4	1157 10	US-09-880-107-2244	Sequence 2244, App
10	1573	99.2	1156 10	US-09-827-854-7	Sequence 7, Appl
11	1570	99.1	1156 10	US-09-827-854-10	Sequence 10, Appl
12	1369.5	86.4	41907 10	US-09-967-013-5	Sequence 5, Appl
13	867.5	54.7	786 10	US-09-925-302-133	Sequence 133, App
14	625	39.4	478 10	US-09-964-824-374	Sequence 374, App
15	625	39.4	478 10	US-09-880-107-2491	Sequence 2491, App
16	463	29.2	356 10	US-09-960-352-5420	Sequence 5420, App
17	458.5	28.9	449 10	US-09-960-352-4726	Sequence 4726, App
18	439.5	27.7	442 10	US-09-960-352-9395	Sequence 9395, App
19	435	27.4	414 10	US-09-960-352-4237	Sequence 4237, App
20	428.5	27.0	416 10	US-09-960-352-8720	Sequence 8720, App
21	422.5	26.7	423 10	US-09-960-352-14047	Sequence 14047, A
22	405.5	25.6	409 10	US-09-960-352-5148	Sequence 5148, App
23	401.5	25.3	416 10	US-09-960-352-9797	Sequence 9797, App
24	401	25.3	253 10	US-09-179-5368-130	Sequence 130, App
25	387.5	24.4	425 10	US-09-960-352-3497	Sequence 3497, App
26	380.5	24.0	390 10	US-09-960-352-1311	Sequence 1311, App
27	379.5	23.9	391 10	US-09-960-352-1278	Sequence 1278, App
28	379.5	23.9	392 10	US-09-960-352-10599	Sequence 10599, A
29	379.5	23.9	401 10	US-09-960-352-4187	Sequence 4187, App
30	375.5	23.7	413 10	US-09-960-352-8042	Sequence 8042, App
31	374	23.6	252 10	US-09-796-116-24	Sequence 24, Appl
32	374	23.6	252 10	US-09-879-341-24	Sequence 24, Appl
33	369.5	23.3	404 10	US-09-960-352-2325	Sequence 2325, App
34	369.5	23.3	405 10	US-09-960-352-14063	Sequence 14063, A
35	351.5	22.2	377 10	US-09-960-352-176	Sequence 176, App
36	350.5	22.1	353 10	US-09-960-352-4914	Sequence 4914, App
37	350.5	22.1	378 10	US-09-960-352-10040	Sequence 10040, A
38	350.5	22.1	386 10	US-09-960-352-10361	Sequence 10361, A
39	348.5	22.0	388 10	US-09-960-352-11986	Sequence 11986, A
40	345.5	21.8	396 10	US-09-960-352-3457	Sequence 3457, App
41	344.5	21.7	377 10	US-09-960-352-10540	Sequence 10540, A
42	338.5	21.4	432 10	US-09-960-352-8934	Sequence 8934, App
43	338	21.3	332 10	US-09-960-352-6552	Sequence 6552, App
44	337.5	21.3	376 10	US-09-960-352-5408	Sequence 5408, App
45	335.5	21.2	373 10	US-09-960-352-3681	Sequence 3681, App

ALIGNMENTS

RESULT 1
US-09-870-759-129
Sequence 129, Application US/09870759
Patent No. US20020177551A1
GENERAL INFORMATION:
APPLICANT: TERMAN, DAVID S
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
FILE REFERENCE: 870759
CURRENT FILING DATE: 2002-01-14
PRIORITY APPLICATION NUMBER: US 60/208,128
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 166
SOFTWARE: PatentIn version 3.1
SEQ ID NO 129
LENGTH: 1156
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (61)..(1014)
OTHER INFORMATION:
US-09-870-759-129
Alignment Scores:
Pred. No.: 4.99e-136
Score: 1585.00
Percent Similarity: 100.00%
Length: 1156
Matches: 317
Conservative: 0


```

US-09-827-854-15 (1-317) x US-09-870-759-129 (1-1156)
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

```

QY	1	MellysValIleuTfPAlaIleuLeuValThrPheLeuAlaGlySGlnAlaVal	20
Db	61	ATGAAGGTTCTGTGGCGCTGCTTGTGTCAATTCTGGCAGATGCCAGGCAAGTG	120
QY	21	GIuGlnAlaValIGluThrGIuProGIuLeuArgGlnInhrGIuTfPGLInSer	40
Db	121	GAGCAAGCGGTGGAGACAGACCAGGAGCCAGACTGCCGCCAGACAGCAAGTGGCAGAGC	180
QY	41	GlyGlnArgTfPGLIleuAlaIleuGlyArgPheTrpAspTfPLeuArgTfPValGlnThr	60
Db	181	GGCCAGCCCTGGAACTGGCACTGGGTGCTTTTGGATTACCTGGCCTGGGTGCACACA	240
QY	61	LeuSerGIuGlnValGlnGlnGluIleuLeuSerSerGlnValThrGlnGluLeuArgAla	80
Db	241	CTGTCTGAGCAGGTGGAGAGAGAGACTGTCTCAAGTCCAGTCAACCCAGAACTAGAGGCG	300
QY	81	LeuMetAspGIuThrMetLysGluLeuLysAlaTfLysSerGIuLeuGlnGluGlnLeu	100
Db	301	CTGATGGACGAGACCATGTAAAGATTGAAGGCTTCAAAATCGAACTGGAGAGAACTAG	360
QY	101	ThrProValIaGIuGluThrArgAlaArgLeuSerLysGIuLeuGlnAlaIaGlnAla	120
Db	361	ACCCCGGTGGCGGAGAGACGCGGGCACGGCTGTCCAAAGNACTGCAAGCGGCGCAAGCC	420
QY	121	ArgLeuGlyAlaAspMetGIuAspValCysGlyArgLeuValGlnTfArgGIuGluVal	140
Db	421	CGCGTGGCGCGGACATGTGAAGAGCTGTGGCGCGCTGGTGCAGTACCGGGGAAGTG	480
QY	141	GlnIlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	481	CAGGCCATGCTGGCCAGACACCGAGAGAGCTGGGGGTGGCTCGCTCCACCTGCGC	540
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	541	AAGCTGGTAAAGCGGCTCTCCGGATGCGCATATCCTGCAGAAAGCCCTGGCAGTGTAC	600
QY	181	GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200
Db	601	CAGCGCGGGGCCCGCAGAGGCGCGAGCGCGGCTTCAGCCCATCCCGAGCGCCTGGGG	660
QY	201	ProLeuValGIuGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro	220
Db	661	CCCCGTGGTGAACAGGGCGCGGTGCGGGCGCCACTGTGGGCTCTCCGTGGCGGCCACCG	720
QY	221	LeuGlnGluArgAlaGlnAlaTfPGLIuArgLeuArgAlaArgMetGIuLeuMetGly	240
Db	721	CTACAGAGAGGGGCCAGAGCCTGGGGGAGACGGCTGCGCCCGCGGATGGAGAGATGGGC	780
QY	241	SerArgThrArgAspArgLeuAspGIuValLysGlnGlnValAlaGluValArgAlaLys	260
Db	781	AGCGGAGCCCGCGACCCCTCGACAGAGGTAAAGACAGCATGTGGCGGAGGTCCGCCCAAG	840
QY	261	LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnIlaPheGlnAlaArgLeuLys	280
Db	841	CTGGAGAGAGCGGCCACAGCAATACGCTCTCAAGGCCAGAGCCTTCCAGAGCGCCCTCAAG	900
QY	281	SerTrpPheGIuProLeuValGluAspMetGlnArgGlnTfPAlaGlyLeuValGluLys	300
Db	901	AGCTGTTTCGAGCCCTGTGGTGAAGACATGCAGGCCCAATGGGCGCTGTGTGGAGAG	960
QY	301	ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis	317
Db	961	GTGCAAGGCTGCGCTGGGCGACACAGCGCGCCCTGTGTGCCAGCAATCATC	1011

```

? Publication No. US2003036057A1
?
? GENERAL INFORMATION:
?
? APPLICANT: Braun, Andreas
? APPLICANT: Bonsel Aruna
? APPLICANT: Kleyen Patrick
? TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH
? TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND THEIR USE
? FILE REFERENCE: 24736-2048
?
? CURRENT APPLICATION NUMBER: US/09/802,640
? CURRENT FILING DATE: 2001-03-09
? NUMBER OF SEQ ID NOS: 122
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 17
? LENGTH: 1156
?
? TYPE: DNA
?
? ORGANISM: Homo sapien
?
? FEATURE:
?
? NAME/KEY: CDS
? LOCATION: (61)...(1014)
? OTHER INFORMATION: Nucleotide Sequence encoding apolipoprotein E
? OTHER INFORMATION: (Apoe)
?
? US-09-802-640-17

```

Alignment Scores:		4.99e-136	Length:	1156
Pred. No.:		1585.00	Matches:	317
Score:		100.00%	Conservative:	0
Percent Similarity:		100.00%	Mismatches:	0
Best Local Similarity:		100.00%	Indels:	0
Query Match:		100.00%	Gaps:	0
DB:		9		
US-09-827-854-15 (1-317) x US-09-802-640-17 (1-1156)				
QY	1	MettysVallelntRpAlaAlaLeuLeuValThrpheLeuAlaGlyGlnAlaIysVal	20	
Db	61	ATGAAGTCTCTGTGGCGCTTGGTCTGTGCATCTCTCGAGAGATCCAGGCAAGTGC	120	
QY	21	GLUGlnAlaValAGlUthrGluProGluProGluLeuAArgGlnGlnIthrGluTPrGlnSer	40	
Db	121	GAGCAAGCGGTGGAGACAGACCAGGACCCAGCTGGCCAGCAGACCGAGTGGCAAGC	180	
QY	41	GlyGlnAArgTPrGluLeuAlaLeuGlyAArgpHeTrpAspIryLeuAArgTPrValGlnThr	60	
Db	181	GGCCAGCGCTGGGAACCTGGCACTGGGTGGTGGATTAACCTGGCGCTGGGTCCACACA	240	
QY	61	LeuSerGlnGlnValGlnGlnGluLeuLeuSerSerGlnValThrnGlnGluLeuArgAla	80	
Db	241	CTGTCTAGAGAGGTGCCAGAGAGAGACTGTCAAGTCCCAAGTCACCCAGAACTAGAGGCG	300	
QY	81	LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu	100	
Db	301	CTGATGGAGCAGACCATGAAAGATTGAAGGCTTAACAAATCGGAACGTGAGAAACAACCTG	360	
QY	101	ThrProValAlaGluGluThrArgAlaAArgLeuSerLysGluLeuGlnAlaIaGlnAla	120	
Db	361	ACCCGGGTGGGGAGAGACCGGGCAAGGCTGTCCAAAGAGCTGCAGGGCGGCAGGCC	420	
QY	121	ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyLysVal	140	
Db	421	CGGCTGGGGCGGACATGAGAGACGTGTGGCGCCGCTGTGTGCATGCCCGGCGAGGTG	480	
QY	141	GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg	160	
Db	481	CAGGCATAGCTCGGCGCAGAGACCGAGAGACTGGGGGTGCGCTCGCTCCCACTGGCG	540	
QY	161	LysLeuAArgLysAArgLeuLeuAArgAspAlaAspAspLeuGlnLysAArgLeuAlaValTyr	180	
Db	541	AAGCTGGCTAAGCGGCTCTCCCGGATGCCAGATACCTTGCAGAAAGCGCTGGAGTGTAC	600	
QY	181	GlnAlaGlyAlaAArgGlnGlyLysArgLysValGlnLysSerLalLeuArgLysArgLeuGly	200	
Db	601	CAGGCGGGGGCGCGAGGGGGCGGACGGGGCTTACGGCCCAATCCGCGAGGCCCTTGGGG	660	

Oy 201 ProLeuValGluGlnGlyArgValArgAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 661 CCCCTGGTGAACAGAGCGCGCTGGCGGCCCACTGTGGGCTCCCTCGCGCGGCAAGCCG 720
 Oy 221 LeuGlnGluArgAlaGlnAlaThrPglGlyLysArgLeuArgAlaGlyMetGluMetGly 240
 Db 721 CTACAGGAGCGGCGCCAGGCTGGGGCGAGCGGCTGGCCGCGGATGAGAGAGATGGGC 780
 Oy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 Db 781 AGCCGACCGCGAGCGCGCTGGACAGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
 Oy 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 Db 841 CTGAG 900
 Oy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 Db 901 AGCTGCTCGAGCGCCCTGGTGAAGACATGCAGCGCCAGTGGCGCGGCTGTGGAGAG 960
 Oy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 961 GTGAGGCTGCGGTGGGACACAGCGCCCTGTGGCCAGCAGCAATCAC 1011

RESULT 3

US-09-827-854-8
 : Sequence 8, Application US/09827854
 : Patent No. US20020123093A1

GENERAL INFORMATION:

: APPLICANT: Zannis, Vassilis
 : APPLICANT: Kyriacos, Kyriakos E.
 : TITLE OF INVENTION: Compounds and methods for lowering
 : TITLE OF INVENTION: cholesterol levels without inducing hypertiglyceridemia
 : FILE REFERENCE: 07180/004003
 : CURRENT APPLICATION NUMBER: US/09/827,854
 : CURRENT FILING DATE: 2001-04-05
 : PRIOR APPLICATION NUMBER: US 09/679,088
 : PRIOR FILING DATE: 2000-10-04
 : PRIOR APPLICATION NUMBER: US 09/544,386
 : PRIOR FILING DATE: 2000-04-06
 : NUMBER OF SEQ ID NOS: 19
 : SOFTWARE: FastSeq for Windows Version 4.0

: SEQ ID NO 8
 : LENGTH: 1156
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : US-09-827-854-8

Alignment Scores:

Pred. No.: 4,99e-136 Length: 1156
 Score: 1585.00 Matches: 317
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-09-827-854-15 (1-317) x US-09-827-854-8 (1-1156)

Oy 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 61 ATGAAGGTTCTGTGGCGCTGTGTCATTCCTGGAGAGATGCCAGGCAAGGTG 120
 Oy 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 Db 121 GAGCAAGCGGTGGAGACAGACCGGAGCCGAGCTGGCGCCAGCAGACCGAATGGCGAGAGC 180
 Oy 41 GlyLysArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLysArgTrpValGlnThr 60
 Db 181 GAGCAGCGCTGGAGAACCTGCGCTTTGGGATTCCTGCGCGTGGCTGCAGAGCA 240
 Oy 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 Db 241 CTGCTGAGCAGGTGCGAGAGAGAGCTGCTCAGCTCCAGGTACCCAGGAACTGAGGCG 300

Oy 81 LeuMetAspGluTrpMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu 100
 Db 301 CTGATGAGCAGAGACCATGAGAGAGTTGAGAGCCCTACAAATGGAGACTGGAGAACTG 360
 Oy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
 Db 361 ACCCGGTGGCGGAGAGAGACCGGCGCACGCTGTCTCAAGAGAGCTGACAGCGCGGAGAG 420
 Oy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
 Db 421 CGGCTGGCGCGGACATGAGAGAGCTGTGCGCGCCGCTGTGGCTGAGTACCGGCGAGAGTG 480
 Oy 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 481 CAGGCCATGCTGGCGCAGAGACCGAGAGAGCTGGGCGGCTTCGCTCCACCTCGCG 540
 Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTrp 180
 Db 541 AAGCTGCGTAAAGCGGCTCTCTCGGAGTCCGATGACCTGCAGAAAGCGCTGTGCAGTAC 600
 Oy 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db 601 CAGCGCGGCGCCCGAGAGCGCGGAGCGCGCTCAAGCCCATCCGAGCGCGCTGGGG 660
 Oy 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 661 CCCCTGGTGAACAGAGCGCGGTGGCGGCCCACTGTGGGCTCCCTGGCGCGGCAAGCCG 720
 Oy 221 LeuGlnGluArgAlaGlnAlaThrPglGlyLysArgLeuArgAlaArgMetGluMetGly 240
 Db 721 CTACAGAGAGCGGCGCCAGCGCTGGGGGAGAGCGGTGGCGCGCGGAGTGAAGAGATGGGC 780
 Oy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 Db 781 AGCCGACCGCGAGCGCGCTGGAGAGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
 Oy 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 Db 841 CTGAG 900
 Oy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 Db 901 AGCTGCTCGAGCGCCCTGGTGAAGACATGCAGCGCCAGTGGCGCGGCTGTGGAGAG 960
 Oy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 961 GTGAGGCTGCGGTGGGACACAGCGCCCTGTGGCCAGCAGCAATCAC 1011

RESULT 4

US-10-044-090-454
 : Sequence 454, Application US/10044090
 : Patent No. US20020137081A1

GENERAL INFORMATION:

: APPLICANT: Olga Bandman
 : TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
 : FILE REFERENCE: PA-0028 US
 : CURRENT APPLICATION NUMBER: US/10/044,090
 : CURRENT FILING DATE: 2002-01-09
 : NUMBER OF SEQ ID NOS: 850
 : SOFTWARE: PERL Program
 : SEQ ID NO 454

: LENGTH: 1291
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: misc feature
 : OTHER INFORMATION: Incyte ID No. US20020137081A1 2514988CB1
 : NAME/KEY: unsure
 : LOCATION: 46
 : OTHER INFORMATION: a, t, c, g, or other
 : US-10-044-090-454


```

|||||
Db 362 ACCCGGTAGCGAGGAGACCGCGGCGCTGTCCAGAGAGCTGCACAGCGCGCGAGGCC 421
Qy 121 ArgLeuGlyAlaAspMetGluAspValGlySerGlyArgLeuValGlnTyrArgGlyVal 140
Db 422 CGGCTGGCGGAGCATGAGACCTGTGGCGCGCTGTGGTCACTACCGCGCGAGGTG 481
Qy 141 GlnAlaMetLeuGlyGlnSerThrGlnLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCGATGCTGGCCACAGACACCGAGAGCTGGGGTGGCTTCCCTCCACCTGGCG 541
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AAGTGCCTAAGCGGCTCTCCGCGATCCGATGCATGCAGAAAGCGCTGGCAGTGTAC 601
Qy 181 GlnAlaGlyAlaArgGlyValGlyAlaGlnArgGlyLeuSerAlaIleArgGlyValGlyLeu 200
Db 602 CAGGCGGGGGCGCGGAGGGCGCGCAGCGCGCTCACCGCCATCCGGAGCGCTGGGG 661
Qy 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 662 CCCCTGGTGGAAACAGGGCGCGGTGGGGCGCCACTGTGGCTCCCTGGCGCGCAGCG 721
Qy 221 LeuGlnGluArgAlaGlnAlaIleArgGlyValArgLeuAlaArgMetGlnGlyMetGly 240
Db 722 CTACAGAGCGCGCGCGCGCTGGGGCGAGCGCGCTGCGCGCGGATGAGAGATGGGC 781
Qy 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
Db 782 AGTCGACCGCGCGACCGCTGTGACAGGTGAAGAGCAGGTGGCGAGAGTGGCGCCAG 841
Qy 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 842 CTGGAGAGAGCAGCGCCACAGATACGCTGCAGCGCGAGGCTTCCAGCGCCCTCAAG 901
Qy 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlyLys 300
Db 902 AGCTGCTTCGAGCGCCCTGTGGAAGACATGCAGCGCAGTGGCGCGCTGGTGGAGAAG 961
Qy 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGCAGGCTGCGGTGGGCACACAGCGCCCGCTGTGGCCAGCAGCAATCAC 1012

RESULT 9
US-09-880-107-2244
: Sequence 2244, Application US/09880107
: Patent No. US20020142981A1
: GENERAL INFORMATION:
: APPLICANT: Horne, Darci T.
: APPLICANT: Vockley, Joseph G.
: APPLICANT: Scherf, Uwe
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
: FILE REFERENCE: 44921-5028-WO
: CURRENT APPLICATION NUMBER: US/09/880,107
: PRIOR FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2244
: LENGTH: 1157
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20020142981A1 M12529
US-09-880-107-2244

```

```

Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.43% Indels: 0
Db: 10 Gaps: 0

US-09-827-854-15 (1-317) x US-09-880-107-2244 (1-1157)
Qy 1 MetLysValLeuTrpAlaAlaLeuLeuValTrpPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAGGTTCCTGTGGGCTGGCTGTGGTACATTCTGGCAGAGATGCCAGGCCAAGGTG 121
Qy 21 GlnGlnAlaValGluThrGlnProGluProGluLeuArgGlnGlnTrpGlnSer 40
Db 122 GAGCAAGGGGTGAGACACAGAGCGGAGCCGAGCTGCGCCACAGACAGCAGAGTGCAGAGC 181
Qy 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpPheTrpPheTrpPheTrpPheTrp 60
Db 182 GCGCAGCGCTGGGAACCTGGCACTGGGTGCTTTGGATTTACCTGGCTGGCTGGCAGACA 241
Qy 61 LeuSerGlnGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db 242 CTGTCTAGACAGGTGCAGAGAGAGCTGCTCAAGTCCCAAGTACCCAGAACTGAGGGCG 301
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
Db 302 CTGATGAGCAGAGACCATGAGAGATTGAAGGCTTCAAAATGGGAACCTGGAGAACACTG 361
Qy 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
Db 362 ACCCGGTAGCGAGAGAACCGGGCGACGGCTGTCCAAAGAGACTGCACAGCGCGCAGGCC 421
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
Db 422 CGGCTGGCGCGGACATGAGAGACTGTGGCGCCCTGTGTCAATACCGCGCGCAGGTG 481
Qy 141 GlnAlaMetLeuGlyGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCGATGCTCGGCGCAGAGCACCGAGAGCTGGGGTGGCGCTGCCCTCCACCTGGCG 541
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AAGTGCCTAAGCGGCTCTCCGCGATCCCGATGCCATGCAGAAAGCGCTGGCAGTGTAC 601
Qy 181 GlnAlaGlyAlaArgGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 200
Db 602 CAGGCGGGGGCGCGGAGGGCGCGCAGCGCGGCTTCCAGCGCGCTTGGGG 661
Qy 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 662 CCCCTGGTGGAAACAGGGCGCGGTGGGGCGCGCACTGTGGCTCCCTGGCGCGCAGCG 721
Qy 221 LeuGlnGluArgAlaGlnAlaIleArgGlyValArgLeuAlaArgMetGlnGlyMetGly 240
Db 722 CTACAGAGCGCGCGCGCGCTGTGGGCGAGCGCGCTGGCGGATGAGAGATGGGC 781
Qy 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
Db 782 AGTCGACCGCGCGACCGCTGTGGAAGACATGCAGCGCAGTGGCGCGCTGGTGGAGAAG 841
Qy 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 842 CTGGAGAGAGCAGCGCCACAGATACGCTGCAGCGCGCAGGCTTCCAGCGCGCTCAAG 901
Qy 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlyLys 300
Db 902 AGCTGCTTCGAGCGCCCTGTGGAAGACATGCAGCGCAGTGGCGCGCTGGTGGAGAAG 961
Qy 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGCAGGCTGCGGTGGGCACACAGCGCCCGCTGTGGCCAGCAGCAATCAC 1012

RESULT 10
US-09-827-854-7

```

Alignment Scores: 3.3e-135 Length: 1157
Pred. No.: 1576.00 Matches: 315


```

Qy 81 LeuMetaspjUthrMetlysgluLeuylsAlaTyrlySsergluLeugluInleu 100
Db 301 CTGATGGAGACCATGAGAGCTTGAAGGCTTCAATGCAATGCACTGGAGAACACTG 360
Qy 101 ThrProValAlaGluGluThrArgAlaArgSerLySgluLeuGlnAlaAla 120
Db 361 ACCCGGGGGGAGAGACGGGGGACGGCTGTCCAGAGAGCTGACGGCGGAGGCC 420
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGly 140
Db 421 CGGCTGGGGCGGACATGAGAGCTGTGGCGCGCTGTGGACATGACCGGAGGTG 480
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCGCATGCTGACCAAGACACCGAGAGCGCGGGTGGCGCTGCCCTCCACTGGCG 540
Qy 161 LysLeuArgLySArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTGCGTAAAGCGGCTCTCCGCGATGCTGACCTGACAGAAAGTGGCTGGCATGTAC 600
Qy 181 GlnAlaGlyAlaArgGluGluValAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 601 CAGGCGGGGGGGCGGAGGGGGCGCGAGCGCGCTCAGCGCATCCGAGCGCGCTGGGG 660
Qy 201 ProLeuValGluGluGlnArgValArgValAlaThrValGlySerLeuAlaGlyInPro 220
Db 661 CCCCTGGTGGAAACAGAGGGCGGGTGGGGCGCCGCACTGTGGCTCTCCGGCGGCAAGCG 720
Qy 221 LeuGlnGluArgAlaGlnAlaThrPcIyGluArgLeuArgAlaArgMetGluGluMetGly 240
Db 721 CTACAGAGAGGGGGCCAGCGCTGGGGGCGAGCGGCTGCGCGCGGATGAGAGATGGGC 780
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 781 AGCGCGAGCCCGGACCGCTGAGAGAGGTGAAGGAGGAGGGCGCGAGAGTGGCGGCCCAAG 840
Qy 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 841 CTGGAGAGAGAGCGCCACAGATACGCTGCAGGGCGGAGGCTTCAGAGCGCGCTCAAG 900
Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 901 AGCTGTTCGAGCCCTGCTGGTGAAGACATGCAGCCCGAGTGGCGGCTGTGGAGAAAG 960
Qy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAspHis 317
Db 961 GTGAGGCTGCGGTGGSCACAGCGCGCCCTGTGCCAGGACAAATCAC 1011

```

RESULT 12

```

US-09-967-013-5
: Sequence 5, Application US/09967013
: Patent No. US20020045840A1
: GENERAL INFORMATION:
: APPLICANT: Stanton, Jr, Vincent P.
: TITLE OF INVENTION: METHOD FOR GENETIC ANALYSIS OF APOE DNA
: FILE REFERENCE: 11926-022001
: CURRENT APPLICATION NUMBER: US/09/967,013
: CURRENT FILING DATE: 2000-10-25
: PRIOR APPLICATION NUMBER: 60/206,613
: NUMBER OF SEQ ID NOS: 91
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 41907
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-967-013-5

```

Alignment Scores:

```

Pred. No.: 1,62e-114 Length: 41907
Score: 1369.50 Matches: 300
Percent Similarity: 60.00% Conservative: 0

```

```

Best Local Similarity: 60.00% Mismatches: 6
Query Match: 86.40% Indels: 194
DB: 10 Gaps: 1
US-09-827-854-15 (1-317) x US-09-967-013-5 (1-41907)
Qy 12 PheLeuAlaGlyCysGlnAlaLysValGluGlnAlaValGluThrGluProGluProGlu 31
Db 20316 TTCCACACAGAGATGCAACRCCAGAGTGGAGAGCGGGTGGAGACAGAGCCGAGCCGAG 20375
Qy 32 LeuArgGlnGlnThrGluTrpGlnSerGlyGlnArgGlyGluLeuAlaLeuGlyArgPhe 51
Db 20376 CTGCCCGACAGACCGAGTGGAGAGCGCGCCAGCGCTGGGAAGTGGCACTGGGTGCTTT 20435
Qy 52 TrpAspTyrIleArgTrpValGlnThrLeuSerGluGlnGluGluGluLeuLeuSer 71
Db 20436 TGGGATTAACCTGCGCTGGGTGACAGACATGTCGTGAGAGGTGACAGAGAGAGTGTCAAGC 20495
Qy 72 SerGlnValThrGluGluLeu----- 78
Db 20496 TCCCAAGTACCCACAGAGAGTGAAGTGAAGTTCGCCATCTGGGCCCTTGACCTCTGGGTG 20555
Qy 78 ----- 78
Db 20556 GAGGCGTATTACCTCCCGAGATCCAGATTTCATTCTGCCCCGTGCGTAACTCTGGGGGG 20615
Qy 78 ----- 78
Db 20616 CTTGGGTCTGCTGGTGTCTAGCTTCCTCCATTTCTGAGCTGCTGGCTTTAGCTCTG 20675
Qy 78 ----- 78
Db 20676 TGGAAATTCCTCTCTACAGCTTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCACT 20735
Qy 78 ----- 78
Db 20736 GCTCTGCTGCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 20795
Qy 78 ----- 78
Db 20796 CTCACGTGTTGCCACAGCGTGGTCTTGAACCTTGGGGCTCAAGCGATCTCCCGGCTCGG 20855
Qy 78 ----- 78
Db 20856 CCTCCCAAGTGTGGGATTAGAGGATGAGCCACCTTGGCGGCTCTAGCTCTCTCT 20915
Qy 78 ----- 78
Db 20916 TCGTCTGTGCTGTGCTGTGATCTGTGATCTGTGATCTGTGATCTGTGATCTGTGATCT 20975
Qy 78 ----- 78
Db 20976 GCTCTGCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 21035
Qy 78 ----- 78
Db 21036 CCCCATCCAGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 21095
Qy 79 -ArgAlaLeuMetaspjUthrMetlysgluLeuylsAlaTyrlySsergluLeuglu 98
Db 21096 CAGGCGGCTGATGAGCGAGACCATGAAGAGTGAAGGCTTCAAAATGGAAGTGAAGAGGA 21155
Qy 98 uGluLeuThrProValAlaGluGluThrArgAlaArgSerLySgluLeuGlnAlaAla 118
Db 21156 ACAACTGACCCCGGTGGGGGAGAGACGGGGGACGGCTGTCAAGAGAGCTGACAGCGCG 21215
Qy 118 aGlnAlaArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGly 138
Db 21216 GCAGGCGCGGCTGGGGCGGAGACATGAGAGAGTGTGGCGCGCTGTGCGAGTACCGCGG 21275
Qy 138 yGluValGlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHis 158
Db 21276 CGAGGTGAGGCGCATGTCTGTGGCGAGAGACCGAGAGAGTGTGGGGTGGCGCTGTGCCCTCCA 21335

```


Oy	158	slenuarqylsleuAargylsArgyleuLeuAargspalaspaspleuGlnlyAargLeuAl	178
Db	21336	CCTGGCAAGCTGyGTAAAGGGGCTCTCCCGCATCCGATGACCTGCAGAAAGCGCTGGC	21395
Oy	178	aValyrglnalagIyAlaAarglucIyAlagluAargglyleuSerAlaIleAargGluAr	198
Db	21396	AGTGTACCAAGCGCGGGGGCCCGGAGGGCCCGACGGCGGGCTTCACGGCCATTCGGGAGCG	21455
Oy	198	gleuGlyProleuValgluGlnGlyArgValAargAlaAarhValglSerLeuAlaGl	218
Db	21456	CCTGGGGCCCCCTGTTGGAAACAGGGCGGGCTGGGGCCCGCACTGTGGGTCTCCCTGGCGCG	21515
Oy	218	ylcInProleuGlnGluAargAlaaglnAlaAtrpylyGluAargLeuAargAlaAargMetGluCl	238
Db	21516	CCAGCGCTATCACAGGAGGGGGCCCAAGGCGCTGGGGGCGAGCGGCTGCAGCGGGATGGAGA	21575
Oy	238	uMetlySerArgTrrArgspArlLeuapsIuVallysgGlnValAlaGlwValAr	258
Db	21576	GATGGGACGCGGACCCCGACCGCTGGACAGGTGAAGAGACAGTGGCGGAGGTGG	21635
Oy	258	gAlaAlyleuGlnGluGlnAlaaglnGlnIleAargLeuGlnAlaGluAlaAphGlnAlaAr	278
Db	21636	CGCCAGCTGGAGAGACAGGCCACACAGATACGCTCGAGCGCCGAGGCGCTTCACAGCCCG	21695
Oy	278	gleuYsserTrrPphGluProleuValGluaspMetGlnAargGlnTrrPalaglyLeuVa	298
Db	21696	CCTCAAGAGCTGTGTGGAGCCCTCTGTGTGAACATGCGAGCCGACGTGGGCGGGCTGGT	21755
Oy	298	lGluIyValGlnAlaAlaValaGlyThrSerAlaAlaProValProSerAspAsnHis	317
Db	21756	GGAGAAAGTGCAGGCTGCGGTGGGACACAGCGCCGCCCTGTGGCCACAGACATTCAC	21813
RESULT 13			
US-09-925-302-133			
; Sequence 133, Application US/09925302			
; Patent No. US20020044941A1			
; GENERAL INFORMATION:			
; APPLICANT: Rosen et al.			
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies			
; FILE REFERENCE: PAL04			
; CURRENT APPLICATION NUMBER: US/09/925.302			
; CURRENT FILING DATE: 2001-08-10			
; PRIOR APPLICATION NUMBER: PCT/US00/05918			
; PRIOR FILING DATE: 2000-03-08			
; PRIOR APPLICATION NUMBER: 60/124,270			
; PRIOR FILING DATE: 1999-03-12			
; NUMBER OF SEQ ID NOS: 896			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 133			
; LENGTH: 786			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-925-302-133			

Alignment Scores:			
Pred. NO.:	7.81e-71	Length:	786
Score:	867.50	Matches:	183
Percent Similarity:	98.39%	Conservative:	0
Best local Similarity:	98.39%	Mismatches:	3
Query Match:	54.73%	Indels:	2
DB:	10	Gaps:	0
US-09-827-854-15 (1-317) x US-09-925-302-133 (1-786)			
QY 132 ArgLeuValGlnTyrArgGlyGluValAlaMetLeuGlyInserThrGluLeu L51			
Db 6 CGCGTGCGCAGTACCGGGCGAGGTGCASGCCATGCTGGCCAGACACCGAGAGCTG 65			
QY 152 ArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArgspAlaSP 171			
Db 66 CGGGTGGCGCTCGCTCCACCTCGCA-CTGGCTAA-CGGCTCCCTCGCGATGCCGAT 123			

QY	172	ASLEUNGILYSARGLEUVALAVALTYRGILNALAGLVALAARGGILUGLVALAGLVALARGGLY	191
Db	124	GACCTGGCAAGACCGCTTGCAAGTGACAGGCCGGGGCCGGCGAGGGGCCGCAAGCGCGC	183
QY	192	LEUSERLALALAEARGGLUARGLEUGLYPROLEUVALGLUGINGLYARGVALARGALALA	211
Db	184	CTCACCGCCATCCGGAGACGCCCTGGGGCCCTGGTGMAACAGGGCCCGCGGGGCCGCC	243
QY	212	THRVALLIGYSERLEUVALAGLYGLINPROLEUNGILUARGALAGLINALATTPGILGLUARG	221
Db	244	ACTGTGGGCTCCCTGGCGCGGACCCGCTACAGAGCGGGGCCACAGGCTTGCGGGAGCCG	303
QY	232	LEUARGALAEARGMETGLUGLUMETGLYSERRTTHRRGASPARGLAUSPGLUVALYS	251
Db	304	CTGCGCGCGCGGATGGAGGAATGGCCAGCCGACCCGCGACCTCTGACGAGGTGAAG	363
QY	252	GLUGLINALAGLVALARGALALYSLDENGILUGLUNGILNALAGLINGILINLEARGLEUGLN	271
Db	364	GAGCAGGTGGCGGAGAGTGGCGGCCCAAGCTGGAGGAGCAGGCCCGCAGATACGCTTCAG	423
QY	272	ALAGLUALAPHEGLINALARGLEULYSERRTPHGEGLUPROLEUVALGLUASPMEGLIN	291
Db	424	GCCGAGGCGCTTCAGAGGCCGCCCTCAAGAGCTGGTTTCAGGCCCTCGGTGGCAAGCAATCAG	483
QY	292	ARGGINTPRALAGLYLEUVALGLILYSAVALGINALALALVALAGLYTHRSERLALALAPRO	311
Db	484	CGCCAGGTGGCGCGGGCTGTGTGAGAGAGGTGACAGGCTCGCGGGGACACAGCGCGGCCCT	543
QY	312	VALPROSERASPARSNHIS	317
Db	544	GTGGCCACGACGACATTCAC	561

```

RESULT 14
US-09-964-824A-374/c
; Sequence 374, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signa
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964, 824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236, 033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 374
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)..(478)
; OTHER INFORMATION: n=a,t,g or c
US-09-964-824A-374

```

Alignment Scores:	
Pred. NO.:	5,45e-49
Score:	625.00
Percent Similarity:	97.63%
Best Local Similarity:	96.92%
Query Match:	39,43%
DB:	10
	Gaps: 0
US-09-827-854-15 (1-317) x US-09-964-824A-374 (1-478)	
QY 1 MetLysValLeuTrpAlaIaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20	

Db 396 ATGAAGTTCTGTGGCTGCTGTGTCACATTCTCGCAGATGCCAGGCAAGGTG 337
Qy 21 Gtuglnalavalglutthrcluproglualeuarglnglnthrcluprglnser 40
Db 336 GAGCAAGCGGTGGAGACAGACCGAGCCNCAGCTGCGCCAGACAGAGAGTGGAGAGC 277
Qy 41 Glylnarqtrpgluleualaleuagllyarppherpasptryleuargtrpvalglnthr 60
Db 276 GGCACGCGCTGGAACTGGCACTGGGTGCTTTGGATTACCTGCGCTGGGTGAGACA 217
Qy 61 Leusergluglnvalglnuglualeu-SerSerglnvalthrnglnuglualeuargl 80
Db 216 CTGCTGTAGCAGGTGCAGAGAGAGCTGCTCCAGCTCCAGAGTCAACCCAGAACTGAGGGC 157
Qy 80 aleumetaspcluthrmetlysglualeuylalatyrllysserglualeuglnuglnle 100
Db 156 GCTGATGACGAGACCATGAAGAGTTGAAGCCCTACAAATCGGAACGAGGAACAACCT 97
Qy 100 uThrProvalalagluginthrarglalaargleuserlysglualeuglnalalaglinal 120
Db 96 GACCCCGGTGGCGGAGAGAGCGCGGCGCTGTCCAGAGAGCTGCAGCGCGCGAGN 37
Qy 120 aargleugllyalaaspmetgluaspyal 129
Db 36 CCGGCTGGCGCGGACATGAGAGAGCTG 9

RESULT 15

US-09-880-107-2491/c
; Sequence 2491, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-NO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2491
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 N33009
; NAME/KEY: unsure
; LOCATION: (1)..(478)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-2491

Alignment Scores:

Pred. No.:	5.45e-49	Length:	478
Score:	625.00	Matches:	126
Percent Similarity:	97.69%	Conservative:	1
Best Local Similarity:	96.92%	Mismatches:	2
Query Match:	39.43%	Indels:	1
DB:	10	Gaps:	0

US-09-827-854-15 (1-317) x US-09-880-107-2491 (1-478)

Qy 1 Metlysalaleutrpalaalaleuvalthrphleualalaglycysglnalalysval 20
Db 396 ATGAAGTTCTGTGGCTGCTGTGTCACATTCTCGCAGATGCCAGGCAAGGTG 337
Qy 21 Gtuglnalavalglutthrcluproglualeuarglnglnthrcluprglnser 40
Db 336 GAGCAAGCGGTGGAGACAGACCGAGCCNCAGCTGCGCCAGACAGAGTGGAGAGC 277

Qy 41 Glylnarqtrpgluleualaleuagllyarppherpasptryleuargtrpvalglnthr 60
Db 276 GGCACGCGCTGGAACTGGCACTGGGTGCTTTGGATTACCTGCGCTGGGTGAGACA 217
Qy 61 Leusergluglnvalglnuglualeu-SerSerglnvalthrnglnuglualeuargl 80
Db 216 CTGCTGTAGCAGGTGCAGAGAGAGCTGCTCCAGCTCCAGAGTCAACCCAGAACTGAGGGC 157
Qy 80 aleumetaspcluthrmetlysglualeuylalatyrllysserglualeuglnuglnle 100
Db 156 GCTGATGACGAGACCATGAAGAGTTGAAGCCCTACAAATCGGAACGAGGAACAACCT 97
Qy 100 uThrProvalalagluginthrarglalaargleuserlysglualeuglnalalaglinal 120
Db 96 GACCCCGGTGGCGGAGAGAGCGCGGCGCTGTCCAGAGAGCTGCAGCGCGCGAGN 37
Qy 120 aargleugllyalaaspmetgluaspyal 129
Db 36 CCGGCTGGCGCGGACATGAGAGAGCTG 9

Search completed: March 15, 2003, 03:03:43
Job time : 80.8753 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:09:47 ; Search time 1289.69 Seconds

(without alignments)
3980.771 Million cell updates/sec

Title: US-09-827-854-15

Perfect score: 1385

Sequence: 1 MKVLMALVTLFLAGCQAKY.....VEKVOAVGTSAAVPSDNH 317

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09827854/ynat_11032003_101610_27495/app_query.fasta.1.3576
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR.SCOR=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pico -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09827854.gcgn.1.1.8826.6runat.11032003_101610_27495 -NCPU=6 -ICPU=3
-NO.XLPT -NO.WMAP -LARGEQUERY -NEG.SCOR=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlm:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	1290.5	81.4	1027	13	BI670367
2	1254	79.1	842	12	BG763371
3	1248	78.7	933	13	BI668318
4	1220	77.0	938	12	BG761746
5	1217.5	76.8	942	13	BI600906
6	1203	75.9	800	13	BM042094
7	1202.5	75.9	922	12	BG472299
8	1199.5	75.7	922	13	BI597743
9	1190.5	75.1	817	12	BG774871
10	1189	75.0	845	13	BI600563
11	1175.5	74.2	845	12	BG829472
12	1152	72.7	790	12	BG707147
13	1149	72.5	919	13	BI551475
14	1139	71.9	757	13	BM042228
15	1129.5	71.3	907	12	BG706129
16	1128	71.2	706	14	BM728696
17	1128	71.0	741	12	BG762924
18	1126	71.0	1110	11	AK010261
19	1124	70.9	804	12	BG702752
20	1123	70.9	834	13	BM042676
21	1120.5	70.7	808	13	BI668329
22	1117	70.5	855	13	BI61362
23	1114	70.3	803	13	BI670350
24	1111	70.1	782	12	BG716776
25	1111	70.1	812	13	BI601551
26	1111	70.1	965	14	BQ677266
27	1108	69.9	757	12	BI603658
28	1105.5	69.7	797	12	BG715366
29	1105	69.7	846	13	BI159757
30	1104	69.7	796	13	BM042153
31	1099	69.3	757	12	BG707750
32	1099	69.3	794	13	BI601279
33	1099	69.3	954	12	BF967543
34	1097	69.2	790	13	BI551066
35	1096	69.1	790	13	BI551811
36	1092.5	68.9	798	12	BG708414
37	1091.5	68.9	891	13	BI549222
38	1085.5	68.5	802	13	BI458355
39	1084	68.4	1100	14	BM914382
40	1083.5	68.4	914	13	BI603523
41	1083	68.3	748	13	BI553065
42	1082.5	68.3	930	12	BF967857
43	1078.5	68.0	812	12	BG769968
44	1078	68.0	688	13	BI544866
45	1077.5	68.0	803	12	BG709360

ALIGNMENTS

RESULT 1
BI670367
LOCUS
DEFINITION B1670367 1027 bp mRNA linear EST 12-SEP-2001
603292738P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5312024 5',
LOCUS
ACCESSION B1670367
VERSION B1670367.1 GI:15584600
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

cdna library arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1AM11790 row: 1 column: 09
High quality sequence stop: 845.

```

FEATURES
source
Location/Qualifiers
1..1027
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5312024"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="RDH10P"
/note="Organ: brain; Vector: pluescript (modified
pbscript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgac
); Oligo-dt primer using primer 5'-gttttttttttttttvtvn-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT
194 a 308 c 400 g 125 t
ORIGIN

```

Alignment Scores:	
Pred. No.:	1,41e-125
Score:	1290.50
Percent Similarity:	90.85%
Best Local Similarity:	89.02%
Query Match:	81.42%
DB:	13
	Gaps:
	2

/note="Organ: skin; Vector: pOMB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."*

BASE COUNT 154 a 250 c 337 g 101 t
ORIGIN

Alignment Scores:
Pred. No.: 7.46e-122 Length: 842
Score: 1254.00 Matches: 260
Percent Similarity: 97.04% Conservative: 2
Best Local Similarity: 96.30% Mismatches: 7
Query Match: 79.12% Indels: 2
DB: 12 Gaps: 0

US-09-827-854-15 (1-317) x BG763371 (1-842)

OY 1 MetLVsValLeuTRPAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
|||||
DB 36 ATGAAGGTTCTGTGGGCTGCTGCTGTGTCACATTCCTGGCAGATGCCAGGCAAGGTG 95
OY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGlnTrpGlnSer 40
|||||
DB 96 GAGCAAGCGGTGGAGACAGAGCGGAGCCGCGGCGCAGACAGACGAGTGGAGAGC 155
OY 41 GlyGlnArgTRPGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
DB 156 GGCCAGCGCTGGGAACCTGGCACTGGGTCCTTTGGGATTAAGTCTGGCGGTGCAGACA 215
OY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
DB 216 CTGTCTGAGCAGAGTCAGAGAGAGCTGCTCAGCTCCAGAGTCAACCGAAGTCAAGGCG 275
OY 81 LeuMetAspGluThrMetTyrGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
|||||
DB 276 CTGATGAGAGACATGAGAGAGTGAAGGCTTCAAAATCGAACTGGAGAGACAACTG 335
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
|||||
DB 336 ACCCGCGGCGGAGAGAGCGGCGGCGCTGCTCAAGAGAGCTGAGCGCGCGAGCGC 395
OY 121 ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
DB 396 CGGCTGGGGCGGACATGAGAGAGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 455
OY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
DB 456 CAGGCCATGCTGGCCAGAGACACGAGAGAGCTGGGGTGGCTGCTGCCAAGCTGGCGC 515
OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
DB 516 AAGCTGCGGTAAGCGGCTCTCCGCGATGCGATGACCTGCAAAAGCGCTGCGCACTGTA 575
OY 181 GlnAlaGlyAlaArgGluGlnGlyValGluArgGlyLeuSerAlaTlaArgGluArgLeuGly 200
|||||
DB 576 CAGGCGGGGGCGCGGAGGCGCGCGAGCGCGCTGAGCGCATCCGAGCGCGCTGGGGG 635
OY 201 ProLeuValGlnGlnGlyValArgValArgAlaAlaThrValGlySerLeuAlaGlyLysPro 220
|||||
DB 636 CCTTG-GTGGAGACAGGCGCGCTGGCGGCGCACTGTGGTCTCCCTGGCGCGCAAGCG 694
OY 221 LeuGlnGluArgAlaGlnAlaAlaTrpGlyLeuArgAlaArgMetGluGlnLysMet-G1 240
|||||
DB 695 CTACAGAGAGCGGCGCGAGCGCTGGCGCGCGATGGAGAGAGATGGCGG 754
OY 240 yserArgTrpArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaVal 260
|||||
DB 755 CAGCGGAGCCCGGAGCGCTGGAGCAGGTGAAGGAGCAGAGTGGCGGAGTGCAGCAAG 814

OY 260 streuGluGlnAlaGlnGlnLeuArg 269
|||||
DB 815 CTGGGAGAGAGCGGCGCGACAGATACGC 842

RESULT 3
LOCUS B1668318 933 bp mRNA linear EST 12-SEP-2001
DEFINITION 603295681P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5314844 5', mRNA sequence.

ACCESSION B1668318
VERSION B1668318.1 GI:15582551
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 933)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki

Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNI)

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LMNI at:
<http://image.llnl.gov>

Plate: L16M11798 row: a column: 21
High quality sequence stop: 780.

FEATURES
source
Location/Qualifiers

1..933
/organism="Homo sapiens"

/db_xref="taxon:9606"
/clone="IMAGE:5314844"

/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"

/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified

pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTT-3',

size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched

for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library

constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 171 a 284 c 359 g 119 t
ORIGIN

Alignment Scores:
Pred. No.: 3.7e-121 Length: 933
Score: 1248.00 Matches: 263

Percent Similarity: 94.08% Conservative: 7
Best Local Similarity: 91.64% Mismatches: 14

Query Match: 78.74% Indels: 4
DB: 13 Gaps: 1

US-09-827-854-15 (1-317) x B1668318 (1-933)

OY 1 MetLVsValLeuTRPAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
|||||
DB 75 ATGAAGGTTCTGTGGGCTGCTGCTGTGTCACATTCCTGGCAGATGCCAGGCAAGGTG 134
OY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGlnTrpGlnSer 40
|||||
DB 135 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTGCGCGCAGACAGACGAGTGGAGAGC 194
OY 41 GlyGlnArgTRPGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||


```

Db 195 GGCACGCGTGGAGACTGGCGACTGGTGGCTTTGGGATTACCTGGCGTGGGTGCAGACA 254
Oy 61 LeuSerGIuGlnValGlnGlnGluLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 255 CTGCTGAGCAGAGGTGAGGAGAGCTGCTGAGCTCCAGGTACACCGAGAACTGAGGGCG 314
Oy 81 LeuMetAspGIuThrMetLysGluLeuLysAlaTyrLysSerGIuLeuGlnGluLeu 100
Db 315 CTGATGGACAGACCAAGAGAGGATTGAAAGCCTTACAAATCGCAACTGGAGGAACAAC 374
Oy 101 ThrProValAlaGlnGluThrArgAlaArgLeuSerLysGIuLeuGlnAlaGlnAla 120
Db 375 ACCCGCGTGGCGGAGAGAGACCGGCGACGGCTGTCCAGAGAGCTGACGGCGCGAGGCC 434
Oy 121 ArgLeuGlnValAspMetGlnAspValLysGlnValArgLeuValGlnThrArgGlnVal 140
Db 435 CGGCTGGGGCGGACATGAGAGAGCTGTGGCGCGCTGTGTCAGTACCGCGCGAGGTG 494
Oy 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuValArgLeuAlaSerHisLeuArg 160
Db 495 CAGGCCATGCTCGCGCAGACAGACCGAGAGCTCGGGTGGCGCTCCACCTGCGCG 554
Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspLeuGlnLysArgLeuAlaValTyr 180
Db 555 AAGCTGCGTAAAGCGGCTCTCCGCGATGCCGATGACCTGCAGAAAGCGCTGGCAGTGTAC 614
Oy 181 GlnAlaGlnValArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
Db 615 CAGGCGGGGGCGCGGAGGCGCGGAGCGCGGCTCAGCGCCATCGCGAGCGCGCTGGGG 674
Oy 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlnSerLeuAlaGlnGln 220
Db 675 CCCCTGGTGAACAGAGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 734
Oy 221 LeuGlnGlnValArgAlaGlnAlaTyrGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 239
Db 735 CTACAGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 794
Oy 239 TglYSerArgThrArgAspArgLeuAsp-GluValLysGlnGlnValAlaGlnValArg 259
Db 795 GGGCAGCGCGAGCGCGCGAGAGGCTGGAAGAGAGTGAAGAGAGAGAGAGAGAGAGAG 854
Oy 259 ValSerLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 279
Db 855 GCCAAGCTGGGGGAGACAGCGCGCGCAATTCGCTTGCAGGCCAA-GCCTTCAGCGCGCG 913
Oy 279 eulYSerTrpPheGlu 284
Db 914 CCAGAGCTGGTTCAAA 930

RESULT 4
LOCUS BG761746 938 bp mRNA linear EST 15-MAY-2001
DEFINITION 602717942F1 NIH_MGC_49 Homo sapiens CDNA clone IMAGE:4841411 5',
ACCESSION BG761746
VERSION BG761746.1 GI:14072399
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 938)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-riemail.nih.gov
Tissue Procurement: ATCC/DC/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be

```

```

FEATURES
    source
        1..938
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4841411"
            /clone_1ib="NIH_MGC_49"
            /tissue_type="melanotic melanoma, high MDR (cell line)"
            /lab_host="PH10B (phage-resistant)"
            /note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
            EcoRI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGCACGAG(G). Size-selected >500bp for average insert size
            1.8kb. Library constructed by Ling Hong in the laboratory
            of Gerald M. Rubin (University of California, Berkeley)
            using 2AP-cDNA synthesis kit (Stratagene) and Superscript
            II RT (Life Technologies). Note: this is a NIH_MGC
            library."

BASE COUNT 171 a 272 c 374 g 121 t
ORIGIN
Alignment Scores:
Pred. No.: 3.32e-118 Length: 938
Score: 1220.00 Matches: 272
Percent Similarity: 92.62% Conservative: 4
Best Local Similarity: 91.28% Mismatches: 11
Query Match: 76.97% Indels: 11
DB: 12 Gaps: 2

US-09-827-854-15 (1-317) x BG761746 (1-938)
Oy 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlnCysGlnAlaLysVal 20
Db 57 ATGAAGATTCTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 116
Oy 21 GlnGlnAlaValAlaGlnThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 40
Db 117 GAGCAAGCGGTGAGAGACAGAGCGCGAGCGCGAGCTCGCGCAGACAGACAGAGAGAG 176
Oy 41 GlyGlnArgTrpGlnLeuAlaLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 60
Db 177 GGGCAGCGCTGGGAGACTGCGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 236
Oy 61 LeuSerGIuGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
Db 237 CTGCTGAGCAGAGGTGAGAGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 296
Oy 81 LeuMetAspGIuThrMetLysGluLeuLysAlaTyrLysSerGIuLeuGlnGluLeu 100
Db 297 CTGATGGACAGACCATGAAAGAGGTTGAAGGCGCTTACAAATCGGAAGTGAAGAACTG 356
Oy 101 ThrProValAlaGlnGluThrArgAlaArgLeuSerLysGIuLeuGlnAlaGlnAla 120
Db 357 ACCCGCGTGGCGGAGAGAGAGCGGGGACGCTGTCCAAAGAGCTGAGCGCGCGAGGCC 416
Oy 121 ArgLeuGlnValAspMetGlnAspValLysGlnValArgLeuValGlnThrArgGlnVal 140
Db 417 CGGCTGGCGCGGACATGAGAGAGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 476
Oy 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
Db 477 CAGGCCATGCTGGCGCAGAGACCGGAGAGCTGCGGGTGGCGCTGCCCTGCCACCTGGCC 536
Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspLeuGlnLysArgLeuAlaValTyr 180
Db 537 AAGCTGCTTAAAGCGGCTCTCCGCGATGACCTCAGAAAGGCGCTGGCAGGTGAC 596
Oy 181 GlnAlaGlnValArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
Db 597 CAGGCGGGGGCGCGGAGGCGCGCGAGCGCTGCTGAGCGCATCGCGAGGCGCTGGGG 656

```

found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LILCM1674 row: c column: 12
 High quality sequence stop: 767.

Location/Qualifiers

KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 800)
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D.

cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LCM1875 row: 1 column: 03
 High quality sequence stop: 792

FEATURES	Location/Qualifiers
source	1. .800

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5420618"
/clone_1id="NH_MGC_112"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pOTR1; site_1: xhoI; site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAGG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NH_MGC Library."

```

BASE COUNT	146 a	239 c	315 g	100 t
ORIGIN				

Alignment Scores:	
Pred. No.:	1,64e-116
Length:	800
Score:	1203.00
Percent Similarity:	99.20%
Best local Similarity:	97.60%
Query Match:	75.90%
DB:	13
Length:	800
Matches:	244
Conservative:	4
Mismatches:	2
Indels:	1
Gaps:	0

US-09-827-854-15 (1-317) x BM042094 (1-800)

OY		1	MeyrsValleuTrpAlaIaleLeuValThrPheueNlaGlyCysGlnAlaIAsVal	20
Db		50	ATGAAGGTCTGTGGCGCTTGCTGTGCACATTCTTCCAGAGATGCCAGGCCAAGGTG	109
OY		21	GluGlnAlaValGIuThrGIuProGILProGIuLeuarGInGlnThrGIuTrpGIuSer	40
Db		110	GAGCAACGGGTGGAGACAGACC GGAGCCGAGCTGGCGCACAGACAGCCAGTGGCAGAGC	169
OY		41	GlyGlnArGrTrpGIuLeuAlaLeuGlyArPheTrpAspTyrLeuArGrTrpValGIuThr	60
Db		170	GGCCAAGCGCTGGAACTGGCACTGGTCGGCTTTGGGATTAACGTGGCGCTGGGTGACAGA	229
OY		61	LeuSerGIuGlnValGIuGInGIuLeuLeuSerSerGlnValThrGInGIuLeuArGAla	80
Db		230	CTGTCTGACGAGTGCAGAGAGAGACTGCTGAGCTCCAGGCTCACCCAGGAACTAGAGGCGC	289
OY		81	LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGInGIuLeu	100
Db		290	CTGATGGACGACCATGGAAGGATTAAAGCCCTACAAATCGGAATCGAGGAACAACATG	349
OY		101	ThrProValAlaGIuGIuThrArgAlaArgLeuSerLysGIuLeuGlnAlaAlaGIuAla	120
Db		350	ACCCCGGTGGCGAGAGAACCGGGGCAAGGCTGTCCAAAGAGACTGCAGCGGGCGCAAGCC	409

QY	121	ArgLeuGluGlyAlaAspMetGluLysPylValGlySerGlyArgLeuValGlnTyrArgGlyGluVal	140
Db	410	CGGCTGGGCGGAGATGAGAGACTGTGGGCCCGCTGTGTCACTACCGCGGAGGCTG	469
QY	141	GlnAlaMetLeuGlyGlnSerThrTluGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	470	CAGGCGATGCTCGGCGCAAGCACCCAGAGAGCTGGCGGTCCCTGCTCCACCTGGCCG	529
QY	161	LysLeuArgTysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	530	AAGCTGCGCTAAAGCGGCTCTCCGCGATGCGATGACGTGCACAAACCGCTGGCAATGAC	589
QY	181	GlnAlaGlyAlaValArgGluGlyAlaGluArgGlyLeuSerAlaIlaArgGluArgLeuGly	200
Db	590	CAGGCGGGGCGCCCGGAGAGGCGCCGAGCGGCGCTTCAGCGCATCGCGAGCGCTGGGG	649
QY	201	ProLeuValGluGlnGlnArgValArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro	220
Db	650	CCCCGTGGTGAACAGGGCGCGCTGGGGCCACACTGTGGCTCTCCCTGGCCGGCCAGCGG	709
QY	221	LeuGlnGluArgAlaGlnAlaAlaThrPylGlyAlaArgLeuArgAlaArgMetGluGlnMetGly	240
Db	710	CTACAGGAGACGGGCCCAAGGCTCTGGGGCGCAGCGGTG-CGCGCGCGGATGAGAGAGATGGCG	768
QY	241	SerArgThrArgAspArgLeuAspGluVal	250
Db	769	ACGCGGACGCGGAGACGCTGTGGAGCAGCGTG	798
RESULT 7	BG472299	927 bp	mRNA
LOCUS	BG472299		linear
DEFINITION	602513830F1 NIH_MGC_16 Homo sapiens CDNA clone IMAGE:4645759 5',		
ACCESSION	BG472299		
VERSION	BG472299.1	GI:13404485	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 927)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabbs-remail.nih.gov		
	Tissue Procurement: ATCC		
	CDNA Library Preparation: Ling Hong/Rubin Laboratory		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov		
	Plate: LCM1419 row: k column: 08		
	High quality sequence stop: 848.		
FEATURES	Location/Qualifiers		
source	1..927		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4645759"		
	/clone_1lib="NIH_MGC_16"		
	/tissue_type="retinodlastoma"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:		
	EcoRI; CDNA made by oligo-dT priming. directionally		
	cloned into EcoRI/XhoI sites using the following 5'		
	adaptor: GGCACGAG(G). Library constructed by Ling Hong		
	in the Laboratory of Gerald M. Rubin (University of		
	California, Berkeley) using Zap-cDNA synthesis kit		
	(Stratagene) and Superscript II RT (Life Technologies).		
	Note: this is a NIH_MGC Library."		
BASE COUNT	170 a 271 c 371 g 115 t		
ORIGIN			

Db 313 CTGATGACAGACCATGGAAGGATTGAAGGCTTACAAATCGGAACCTGGAGAACACTG 312
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 373 ACCCGGCTGGGAGAGACGCGGACCGCTGTCCAAAGAGCTGCAGGCGGCGAGGCC 432
OY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 433 CGGCTGGGCGGACATGAGGAGCGTGTGGCGCGCTGCTGTCAGTACCGCGCGAGGTC 492
OY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 493 CAGGCCATCTGCTGGCCAGAGACCGAGAGCTGCGGGTCTGCTCCCTCCACCTGCGC 552
OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 553 AACCTCGTGAAGGCTCTCTCCGCGATGCGGATGACCTGCAGAGGCGCTGCGAGTTCAC 612
OY 181 GlnAlaGlyAlaArgGluGlnGlnAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 613 CAGCGCGGCGCGCGGAGGCGCGCGAGCGGCTC-AGCGCCATCCGCGAGCGCTGGGG 671
OY 201 ProLeuVal-GluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGln-P 220
Db 672 CCCCTGTGGAGACAGGCGGGGTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 731
OY 220 LeuGlnGlnArg-AlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGlnMet 239
Db 732 CGCTACAGAGCGGCGCGCGCGCTGCGCGAGCGGATACCGCGCGCGAGTGAAGACAG 791
OY 240 GlySerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAla 259
Db 792 GGGAGCGGACCGGACCGCGCTGAGAGGTGAAGAGACAGCGTGGCGGGGTGGGGCCA 851
OY 260 LysLeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleArgGln 276
Db 852 CAACTGAAGAAACAGCCCAAGAAATACGCTTA--GCCGAAGGCTTCAGG 899

RESULT 9
Bg774871 817 bp mRNA linear EST 15-MAY-2001
LOCUS Bg774871 602649975F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4760827 5',
DEFINITION mRNA sequence.
ACCESSION Bg774871
VERSION Bg774871.1 GI:14045188
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 817)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LRCM1612 row: e column: 20
High quality sequence stop: 813.
Location/Qualifiers
1. 817
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4760827"
/clone_1db="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"

/note="Organ: prostate; Vector: pOTB7; Site_1: xhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC Library."

BASE COUNT 148 a 242 c 324 g 103 t

ORIGIN

Alignment Scores:
Pred. No.: 3.5e-115 Length: 817
Score: 1190.50 Matches: 252
Percent Similarity: 97.32% Conservative: 2
Best Local Similarity: 96.55% Mismatches: 5
Query Match: 75.11% Indels: 4
DB: 12 Gaps: 0

US-09-827-854-15 (1-317) x Bg774871 (1-817)

OY 1 MetLysValLeuThrPalAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 40 ARGAAAGTTCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 99
OY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 100 GAGCAAGCGGTGGAGACAGACAGCGGAGCGGAGCTGCGCCAGACAGACAGCGAGTGCGAGAC 159
OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTyrPalGlnThr 60
Db 160 GGCACACGCTGGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 219
OY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db 220 CTGTCTGACAGAGTGCAGAGAGAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 279
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
Db 280 CTGATGAGACAGACCATGAAAGGATTGAAGGCTTACAAATCGGAATCGAGAGAACACTG 339
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 340 ACCCGGCTGGGAGAGAGCGGCGGACGCTGTCTCCAAAGACTCTCAGCGCGGCGAGGCC 399
OY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 400 CGGCTGGCGCGGACATGAGAGACGTGTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 459
OY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 460 CAGGCCATCTGCGCGAGAGACCGGAGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 519
OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 520 AAGCTCGTGAAGCGGCTCTCCGCGAGTGCAGTCACTCAAGAGGCTGCGCAAGTTCAC 579
OY 181 GlnAlaGlyAlaArgGluGlnGlnAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 580 CAGCGCGGCGCGCGGAGGCGCGCGAGCGGCTGAGCGCATCGCGAGCGCGCGCTGGGG 639
OY 201 ProLeuValGluGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
Db 640 CCCCTGTGGAAGAGGCGCGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 698
OY 221 LeuGlnGlnArgAlaGlnAla--TrpGlyGluArgLeuArgAlaArgMetGluGlnMetG 240
Db 699 CTACAGAGACCGGCGCGAGCGCTTGGGGCGAGCGGATGCGCGCGGAGAGGAGATGG 758
OY 240 LysSerArgThrArgAspArgLeuAspGluValLysGluGlnGlnAlaGluValArgAla 259
Db 759 GCCACCGGAGCCG--GAACTGCTGAGAGAGTGAAGAGACAGGTGGCGGAGGTGCGCGCC 815

QY	61	LeuSerGIuGIInValGIInGIuLeuLeuSerSerGIInValThGIInGIuLeuAryala	80
QY	61	LeuSerGIuGIInValGIInGIuLeuLeuSerSerGIInValThGIInGIuLeuAryala	80
Db	253	CGTGTGACACAGTGCACAGAGAGCGTGTCTACACTCCACAGCTCACCCAGAACATGAGGCG	312
QY	81	LeuMetAsGIuThMetLySGIuLeuLysAlaTyrLySGIuLeuGIuInGIuInGIu	100
Db	313	CTGATGACAGACCATGAAGAGTTGAAGGCTTACAAATTCGAACTGGAGGAACAACCTG	372
QY	101	ThrProValAlaGIuGIuThrArgIlaArgLeuSerLySGIuLeuGIAlaAlaGIInAla	120
Db	373	ACCCGGGTGGCGAGAGACCGCGGCACGGCTGTCCAGAGACTCGAGCGGCGCAGGCC	432
QY	121	ArgLeuGIuAlaAspMetGIuAspValCysGIuArgLeuValGIuTyrArgGIuVal	140
Db	433	CGGCTGGGGCGGACATGAGAGACTGTGTGGCGCCCTGTGTGACATACCGCGGAGGTG	492
QY	141	GIuAlaMetLeuGIuGIuSerThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg	160
Db	493	CAGGCAATCTCTGGCCAGACACCAAGAGAGTGGGGTGTGTGCTCCCTCCACTGTGGC	552
QY	161	LysLeuArgLyArgLeuLeuArgAspAlaAspAspLeuGIuLysArgLeuAlaValTyr	180
Db	553	AAGCTGCGTACACGGCTCTCCGCATGTCCGTACCTTCACAGAACGCCCTGGCACTGTAC	612
QY	181	GIuAlaGIuAlaArgGIuGIuValArgIuArgLyLeuSerAlaIleArgGIuArgLeuGIu	200
Db	613	CAGGCGGGGGCCCGGAGGGCGCCAGGGCGGCTTCAGGCGCATCGGAGGCGCTGGGG	672
QY	201	ProLeuValGIuGIuGIuArgValArgAlaAlaThrValGlySerLeuAlaGIuGIuPro	220
Db	673	CCCTGTGTGTGAACAGGGCGCGTGTGGCGCCGACCTGTGGCTCCCTGGCCAGCCGCG	732
QY	221	LeuGIuGIuArgAlaGIuAlaThrGIuGIuArgLeuArgAlaArgMetGIuGIuMet	239
Db	733	CTACAGAGACGGGGCCACAGCGCTGTGGGCGAGCGGTG-CGCGGCGGATGAGAGATG	788
RESULT	13		
LOCUS	B1551475		
DEFINITION	6031943141 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5274003 5',		
ACCESSION	B1551475		
VERSION	B1551475.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 919)		
JOURNAL	NIH-MGC http://mgi.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgaabs-remail.nih.gov		
	Tissue Procurement: Miklos Palkovits, M.D., Ph.D.		
	cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shizaki		
	Toshiyuki and Piero Carninci (RIKEN)		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNLN at:		
	http://image.lnl.gov		
	Plate: L1A11691 row: 1 column: 04		
	High quality sequence scop: 812.		
FEATURES	Location/Qualifiers		
source	1..919		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5274003"		
	/clone_lib="NIH_MGC_95"		
	/tissue_type="hippocampus"		
	/lab_host="DH10B"		

DB	Query	Alignment Scores:	Score:	Percent Similarity:	Best local Similarity:	Query Match:	DB:
US-09-827-854-15	(1-317) x B1551475 (1-919)	9.75e-111	1149.00	93.63%	92.51%	72.49%	13
DB	1	MetlyValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	247	3	16	4	0
DB	73	ATGAAGATTCTGTGGCTGGCTGGTGGTGCATTTCTCGAGGATGCGACAGCCAGGTG	132	3	16	4	0
DB	21	GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer	40	3	16	4	0
DB	133	GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGGCCAGACAGACCGAATGGCAGAGC	192	3	16	4	0
DB	41	GlyGlnArgTrpGluLeuAlaLeuGlyLysArgPheTrpAspTrpLeuArgTrpValGlnThr	60	3	16	4	0
DB	193	GCCAGCGCTGGAGACTGGCACTGGTGGCTTTGGATTAACGTGGCTGGGTGGAGACA	252	3	16	4	0
DB	61	LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80	3	16	4	0
DB	253	CTGTCTGAGCAGGTCTCAGAGAGAGTGTCTCAAGCTCCAGGTCCAGACAGAGTGAAGCGG	312	3	16	4	0
DB	81	LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGluLeu	100	3	16	4	0
DB	313	CTGAATGAGACAGACCATGAGAGATTGAAGGCTTCAATGCAATGCAATGAGACACTG	372	3	16	4	0
DB	101	ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120	3	16	4	0
DB	373	ACCCGGTGGCGGAGAGACCGGGGACGGGCTGCTCAAGAGAGCTGACAGCGGCGAGGCC	432	3	16	4	0
DB	121	ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyVal	140	3	16	4	0
DB	433	CGGCTGGGGCGCAGATGAGAGACTGTGGCGGCCCTGTGTGCACTGACCGCGGAGGTG	492	3	16	4	0
DB	141	GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg	160	3	16	4	0
DB	493	CAGGCCATGCTCGGCAAGACACCGAGAGAGCTGGGGTGGCTTCGCTCCACTGGCGC	552	3	16	4	0
DB	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180	3	16	4	0
DB	553	AAGTCGCGTAACCGGCTCTCCGCGATCCCGATGACCTGAGAAAGCGCTGCGAGTAC	612	3	16	4	0
DB	181	GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200	3	16	4	0
DB	613	CAGGCGGGGGCGGAGTGGCGCCAGGCGCGTGTACACGCGCATCCGCGAGCGCTGGGG	672	3	16	4	0
DB	201	ProLeuValGluGlnGlyArgValAlaArgAlaThrValGlySerLeuAlaGlyGlnPro	220	3	16	4	0
DB	673	CCCTGGTGGGAACAGGGCGCGTGGGGCGCCATGTGGGCTCCCTGGGCGAGCCCT	732	3	16	4	0
DB	221	LeuGlnAlaArgAlaGlnAlaAlaTrpGlyLysLysLeuArgAlaArgMetGluLysMetGlu	240	3	16	4	0
DB	733	A--CAGGAGCGGGCGGACAGGCTGGGGCGAGGGGTGGCGCGCGGATGAGAGAAATGGG	790	3	16	4	0
DB	240	YSerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys	260	3	16	4	0
DB	791	CAGCGGAGCCCGGAGCG--CTGGACAGAGTAAAGAGCACTGGGGCGGGGG--TGGCGCAA	847	3	16	4	0


```

QY      260  sLeuGlUgInGlnAlaGln 266
Db      848  GCTGAGAGACCAGCAGAG 866

RESULT 14
LOCUS    BM042228              757 bp    mRNA    linear    EST 07-NOV-2001
DEFINITION 60361618F1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5557004 5',
            mRNA sequence.
ACCESSION  BM042228
VERSION    BM042228.1 GI:16771495
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE  1 (bases 1 to 757)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cga@bbs-remail.nih.gov
            Tissue Procurement: DCTD/DRP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2021 row: k column: 21
            High quality sequence stop: 757.

FEATURES
    source
        1..757
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5557004"
            /clone_lib="NIH_MGC_112"
            /tissue_type="melanotic melanoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: skin; Vector: pOT87; Site_1: XhoI; Site_2:
            EcoRI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGCACAGAG(G). Library constructed by Ling Hong in the
            Laboratory of Gerald M. Rubin (University of California,
            Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
            Superscript II RT (Life Technologies). Note: this is a
            NIH-MGC library."
BASE COUNT 134 a 229 c 236 g 98 t
ORIGIN
Alignment Scores:
Pred. No.:      8.38e-110      Length:      757
Score:          1139.00        Matches:    235
Percent Similarity: 99.16%      Conservative: 0
Best Local Similarity: 99.16%      Mismatches: 1
Query Match:    71.86%         Indels:     2
DB:             13            Gaps:        0

US-09-827-854-15 (1-317) x BM042228 (1-757)

QY      1  MetLysValLeuTrpAlaAlaLeuLeuValTrpPheLeuAlaGlyCysGlnAlaLysVal 20
Db      50  ATGGAAGTTCTGTGGGCTCGTGTGCTGCATCTCTGCGAGGATGCGAGGCCAGAGTG 109

QY      21  GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGluGlnThrGluTrpGlnSer 40
Db      110  GAGCAAGCGGTGAGACACAGCCGAGCGGAGCTCGCCAGACAGCCAGTGGCAGAGC 169

QY      41  GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db      170  GGCACGCGCTGGGAACCTGGCACTGGCTCTTTGGGATTACCTGGCTGGGTGCAGACA 229

```

```

QY      61  LeuSerCluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db      230  CTGCTGTGAGCAGGTGACAGGAGAGCTGCTCAGCTCCAGCTCACAGCAACAGAGAGCGC 289

QY      81  LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
Db      290  CTGATGACAGACCATGAAAGAGTTGMAAGCCTTACAAATCGGAACACTGAGAGAACTACTG 349

QY      101  ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db      350  ACCCGGTGCGAGAGAGACCGCGGCGCTGTCCAAAGAGACTGCAGGCGCGCAGAGCC 409

QY      121  ArgLeuGlnLysAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db      410  CGGCTGGCGCGGACATGAGGACGTGTGGCGCGCTGTGCATACCCCGCGGAGGTG 469

QY      141  GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuLysSerHisLeuArg 160
Db      470  CAGGCCATGCTCGGCCACAGACACCGAGAGCTCGGGGTGCGCTTCCCTCCACTGCGC 529

QY      161  LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db      530  AAGCTCGTAAAGCGGCTCTCCGCGATGCCGATGACCTGCAGAAAGCGCTGCGCAGTGTAC 589

QY      181  GlnAlaGlyAlaArgGluGlnGluAlaGluArgGlyLeuSerAlaAlaLeuArgGluGln 200
Db      590  CAGGCGGGGCGCGCAGAGGCGCGGAGCGGCGCTCAGCGCCATCCGCGCAGCGCCCTGGG 649

QY      201  ProLeuValGluGlnGlnLysArgVal-ArgAlaAlaThrValGlySerLeuAlaGlySer 220
Db      650  CCCCTGTGTGAACAGGCGCGCGCTTGGGGCGCGCACTGTGTGCTTCCCTGCGCGGCAAGC 709

QY      220  OleuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuAlaArgMet 236
Db      710  GCTACAGAGACGGGCG-CAGGCGTGGGGCGAGCGGCGCGCGGATG 757

RESULT 15
LOCUS    BG706129              907 bp    mRNA    linear    EST 07-MAY-2001
DEFINITION 60266903F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4792030 5',
            mRNA sequence.
ACCESSION  BG706129
VERSION    BG706129.1 GI:13981169
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE  1 (bases 1 to 907)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cga@bbs-remail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shlraxi
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10669 row: 1 column: 23
            High quality sequence stop: 832.

FEATURES
    source
        1..907
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4792030"
            /clone_lib="NIH_MGC_96"
            /tissue_type="hypothalamus"
            /lab_host="DH10B"
            /note="Organ: brain; Vector: pBluescriptPR (modified)

```


Bluescript KS+; Site 1: BamHI; Site 2: SalI-XhoI (gtcagag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 171 a 260 c 358 g 117 t 1 others

Alignment Scores:

Pred. No.:	1.08e-108	Length:	907
Score:	1129.50	Matches:	244
Percent Similarity:	94.62%	Conservative:	2
Best Local Similarity:	93.85%	Mismatches:	10
Query Match:	71.26%	Indels:	6
DB:	12	Gaps:	2

US-09-827-854-15 (1-317) x BG706129 (1-907)

```
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 75 ATGAAGGTTCTGTGGGCTGCTGGTGGTGCATTCCTGGCAGATGCCAGGCCAAGGTG 134
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 135 GAGCAGACCGGTGGAGACAGAGCCGAGCTGGCCAGACAGACCAGATGCCAGGTG 194
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTyrPylGlnThr 60
DB 195 GGCCAGCGCTGGGAACTGGCACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 254
QY 61 LeuSerGluGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 255 CTGTCTGACAGAGTGCAGAGAGAGCTGCTCAGCTCCAGGTCACCCAGAACCTGAGGCG 314
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
DB 315 CTGATGAGACAGACCATGACAGAGAGTGGAGGCTTACAAATCGAATCGAGAGACACTG 374
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGlu-LeuGlnAlaAlaGlnAl 120
DB 375 ACCCGGTGGCGAGAGAGAGCGGGCAGCGCTGCCAAGAGAGTGCAGGCGCGCAGGC 434
QY 120 ArgGluGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 435 CCGGCTGGCGCGGACATGAGAGAGCTGCGCGCCCTGGTGCAGTACCAGCGCGAGGT 494
QY 140 LcGlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuAr 160
DB 495 GCAGGCCATGCTGCGCCAGACAGACAGAGAGTGGCGGTGGCTGCCCTCCACTGCG 554
QY 160 GlySerLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 555 CAAGCTGCGTAAAGCGGCTCCCGGATGCGCATGACCTGCAGAAAGCGCTGCGAGTGA 614
QY 180 TglnAlaGlyAlaArgGluGlyAlaGlnArgLysLeuSerAlaAlaArgGluArgLeu-G 200
DB 615 CCAGGCCCGGCGCGCGAGGCGCGAGCGCGCTCAGCCCATCCGAGCGCGCTGGG 674
QY 200 LysProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnP 220
DB 675 GCCCCTGGTGGAGACAGAGCGCGGTGGCGCGCATCTGGGCTCCCTGGGCGCGCAGC 734
QY 220 ToleuGlnGluArgAlaGlnAlaTrpGlyGlyArgLeuArgAlaArgMetGluGlnMetG 240
DB 735 CGCTACAGAGAGCGGCGCAGGCTG--GGGCGAGCGGTGGCGCGCGATGAGGAGATGG 791
QY 240 LysArgTrpThrArgAspArgLeuAspGluValLysGlnValAlaGluValArg 258
DB 792 GCAG--CGGACCGCG--ACGCTGACAGAGTG--AAGGAGCAGTGGGGAAGTGTGCCG 842
```

Search completed: March 14, 2003, 20:13:52
Job time : 1294.69 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:08:57 ; Search time 1477.5 Seconds
(without alignments)
3998.551 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_203
Perfect score: 1014
Sequence: 1 MKVLMALVLTVFLGCOAKV.....AREGAEGLSAIRERLPLV 203

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+P2n.model -DEV=xlp
-O=/cgn2.1/USPTO/spool/US09827854/lynat_11032003_101610_27486/app_query.fasta_1.3576
-DB=gen5mb1 -QPMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTMT=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09827854 -GCN.1_1_13008 -etunat_11032003_101610_27486 -NCPU=6 -ICPU=3
-NO_XLPRX -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rnd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1014	100.0	1110	6 E00359	E00359 cDNA coding
2	1014	100.0	1110	6 E00823	E00823 DNA sequenc
3	1014	100.0	1147	6 AX302545	AX302545 Sequence
4	1014	100.0	1156	6 BD004278	BD004278 Apo E hum
5	1014	100.0	1156	6 HUMAPOE3	K00396 Homo sapien
6	1014	100.0	1186	6 BC003557	BC003557 Homo sapi
7	1007	99.3	1110	6 E08423	E08423 DNA coding
8	1006	99.2	1156	6 BD004277	BD004277 Apo E hum
9	1005	99.1	1157	6 AX333278	AX333278 Sequence
10	1005	99.1	1157	6 AX405597	AX405597 Sequence
11	1005	99.1	1157	6 H15975	H15975 Sequence 1
12	1005	99.1	1157	6 HUMAPOE	M12529 Human apoli
13	1002	98.8	1156	6 BD004279	BD004279 Apo E hum
14	984	97.0	660	6 A62340	A62340 Sequence 28
15	984	97.0	660	6 A62342	A62342 Sequence 30
16	984	97.0	660	6 A62344	A62344 Sequence 32
17	984	97.0	660	6 AR075563	AR075563 Sequence
18	920	90.7	1178	9 MFAPOE	X13887 Monkey mRNA
19	825.5	81.4	5491	9 AF261279	AF261279 Homo sapi
20	825.5	81.4	41907	6 AX358722	AX358722 Sequence
21	825.5	81.4	41907	9 AF050154	AF050154 Homo sapi
22	825.5	81.4	107567	9 AC011481	AC011481 Homo sapi
23	808	79.7	5515	9 HUMAPOE4	M10065 Human apoli
24	789	77.8	1138	4 AF303830	AF303830 Tupai gi
25	782	77.1	5413	9 AF261280	AF261280 Pan trogl
26	769	75.8	1045	10 MUSAPOE	M12414 Mouse apoli
27	769	75.8	1104	10 BC028816	BC028816 Mus muscu
28	760.5	75.0	965	6 AX384545	AX384545 Sequence
29	760.5	75.0	1108	4 BTRAPOLPE	X61171 B. taurus mR
30	760.5	75.0	5617	6 AX384541	AX384541 Sequence
31	760.5	75.0	6026	6 AX384539	AX384539 Sequence
32	749	73.9	208239	2 AC021988	AC021988 Homo sapi
33	742.5	73.2	1154	4 BRABOEMR	X64839 B. taurus mR
34	735.5	72.5	1060	4 RABAPOLP	M36603 Rabbit apol
35	730	72.0	4762	6 BABAPOE	M29322 Baboon apol
36	724.5	71.4	1126	6 ARI64342	ARI64342 Sequence
37	724.5	71.4	1126	6 AR205885	AR205885 Sequence
38	718	70.8	951	10 S76779	S76779 RAPOE-apol1
39	718	70.8	1069	10 RATAPOE	J00705 rat apoli
40	715.5	70.6	1122	4 SRAPOE	X72835 S. scrofa mR
41	692	68.2	959	10 MUSAPOE	M73480 Mus musculu
42	675	66.6	528	9 HSAPOE	X00199 Human mRNA
43	649	64.0	228698	2 AC127479	AC127479 Mus muscu
44	649	64.0	237653	2 AC073760	AC073760 Mus muscu
45	625	61.6	478	6 AX330507	AX330507 Sequence

RESULT 1

ALIGNMENTS

LOCUS	E00359				
DEFINITION	cDNA coding human apolipoprotein E3.				
ACCESSION	E00359				
VERSION	E00359.1	GI:2168646			
KEYWORDS	JP 1985118189-A/1.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 1110)				
TITLE	Teranishi, Y., Takamatsu, N., Matsui, Y., Kimura, M. and Ikeda, Y.				
JOURNAL	DNA FRAGMENT				
COMMENT	Patent: JP 1985118189-A 1 25-JUN-1985;				
	MITSUBISHI CHEM IND LTD				
	OS	human	06510189		

PN		JP	196818183-A/1
PD		ZP	29-NOV-1985
PE		JF	29-NOV-1983 JP 1983224980
PI			TEANISHI YUTAKA, TAKAMATSU NOBUHIRO, MATSUI YASUSHI, PI
			KIMURA MASAO,
PI			IKEDA YASUKO
PC			C12N15/00.C07H21/04//C12P21/00;
CC			strandedness: Double;
CC			topology: Linear;
CC			hypothetical: No;
CC			anti-sense: No;
CC			*source: tissue_type=Liver;
FH			Location/Qualifiers
FH			
FT			15..968
FT			/product='apolipoprotein E3 precursor' FT
sig_peptide			15..68 /product='apolipoprotein E3 signal peptide'
mat_peptide			69..965 /product='apolipoprotein E3'.
FT			

FEATURES	location/Qualifiers
source	1. 1110 /organism="Homo sapiens" /db_xref="taxon:9606"
BASE COUNT	198 a 353 c 416 g 143 t
ORIGIN	

Alignment Scores:	
Pred. No.:	1,53e-71
Score:	1014.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	6
Length:	1110
Matches:	203
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-827-854-15_COPY_1-203 (1-203) X E00359 (1-1110)

QY	1	MeltyValLeuTPAlaAlaLeuValThPheuAlaGlySGAlaIaVal	20
Db	15	ATGAAGGTTCGTGGCGTCGCTCTGTGCATTCCTGGCAGGATCCAGGCCAAGTC	74
QY	21	GIuGIAlaIaValGIuThrGIuProGIuProGIuLeuArgInGIuThrGIuTPGIuSer	40
Db	75	GAGCAACCGGTGGAGACAGAGCCGAGCCGAGCTGGCCAGCAGACCGAGTGGCAAGC	134
QY	41	GIyGIAlaTrpPAlaLeuAlaLeuGIyArgPheTrpAspIyIleuArgTPAlaGIuThr	60
Db	135	GGCCAGCGCTGGAGACTGGCACTGGTGGCTTTTGGGATTAACCTGGCCCTGGGTCAACACA	194
QY	61	LeuSerGIuGIuAlaGIuGIuGIuLeuLeuSerSerGIuAlaThrGIuGIuLeuArgAla	80
Db	195	CTGTCTAGCAGGTGGCAGGAGAGACTGCTCAAGCTCCACAGGTCAACCCAGGAACGTAGGGCGC	254
QY	81	LeuLeuTrpGIuTrpMetIySGIuLeuIyAlaIyTrpIySerGIuGIuGIuGIuGIuGIuLeu	100
Db	255	CTGTGAGACGAGCCATGAAGAGATTAAGAGCCCTACAAATGTGGAACTGGAGGAACAACCTG	314

QY	101	ThpProValAlaIugluThrArgAlaArgLeuSerLysGluGlnAlaIaAlaGlnAla	120
Db	315	ACCCCGGTGGGGAGAGACACCCGGGGCACGGCTGTCCAAAGACCTGCAGGGGGCCAGGCC	374
QY	121	ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal	140
Db	375	CGGCTGGGGCGGGAGACATGGAGAGAGCTGTGGCGCGCTGTGGTCAGTACCGGGGAGAGTG	434
QY	141	GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	435	CAGGCGATGCTCGGCGCAGAGACACCGAGAGAGCTGGGGGTGGCGCTCGGCTCCACCTGC	494
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	495	AAGCTGCCTAAGCCGGCTCTCCCGGATCCGATACCTGCAGAAAGCCCTCGGCACGTCTAC	554
QY	181	GlnAlaGlyAlaArgGluGluAlaGluArgGlyLeuSerLaiLeArgLValArgLeuGly	200
Db	555	CAGGCGGGGGCCCGCGAGGGGGCCGAGAGCGGGCTCAGCGCCATCCGCGAGCGCGTGGGG	614
QY	201	ProLeuVal 203	
Db	615	CCCCCGGTG 623	

RESULT 2			
E00823	1110 bp	DNA	linear
LOCUS			PAT 29-SEP-1997
DEFINITION	DNA sequence coding for human apolipoprotein E and its signal		

ACCESSION E00823
VERSION E00823.1 GI:2169084
KEYWORDS JP 198609697-A/1.
SOURCE unidentified.
ORGANISM unidentified

REFERENCE	1 (bases 1 to 1110)
AUTHORS	Teraishi, Y., Matsui, Y., Ikeda, Y. and Kimura, M.
TITLE	PRODUCTION OF HUMAN APOLIPOPROTEIN E-LIKE PROTEIN
JOURNAL	Patent: JP 1986096997-A 1 15-MAY-1986;
COMMENT	MITSUBISHI CHEM IND LTD
OS	Human (Homo sapiens)

PN		JP 19860969/-A/1			
PD		15-MAY-1986			
PF		16-OCT-1984 JP 1984216987			
PI		TERANISHII YUTAKA, MATSUI YASUSHI, IKEDA YASUOKO, KIMURA MASAKO			
PC		C12P21/00,A6IK35/74,A6IK37/04,C12N15/00,(C12P21/00,C12R1:19),			
PC		(C12N15/00,			
PC		C12R1:19):			
CC		strandedness: Double;			
CC		topology: linear;			
CC		hypothetical: No;			
CC		anti-sense: No;			
CC		*source: tissue=Liver;			
FH		key Location/Qualifiers			
FT		3'UTR	1..14		
FT		sig-peptide	15..68		
FT		/product='human apolipoprotein E signal peptide			
FT		peptide	69..968		
FT		/product='human apolipoprotein E'			
FT		FT	3'UTR		

FEATURES	969.	1110.	Location/Qualifiers
source	1.	1110	/organism="unidentified" /db_xref="taxon:32644"
BASE COUNT	198 a	353 c	416 g 143 t
ORIGIN			
Alignment Scores:			
Pred. No.:	1 53e-71	Length:	1110
Score:	1014.00	Matches:	203
Percent Similarity:	100.00%	Conservative:	0

Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 6
US-09-827-854-15_COPY_1_203 (1-203) x E00823 (1-1110)
Mismatches: 0
Indels: 0
Gaps: 0

OY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
15 ATGAAGGTTCTGTGGGCTGCGCTGTGGTGCATCTCCGAGATCCAGATGCCAGGCCAAGGTG 74
OY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
75 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGGCCAGACAGACGAGTGGAGAGC 134
OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
135 GGGCAGCCGCTGGGAACCTGGCACTGGCTCTTTGGGATTACCTCGCTGGGTGCAGACA 194
OY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
195 CTGCTGAGCAGAGTCAGAGAGAGCTGCTCAGCTCCAGAGTCACCCAGGAACCTGAGGGCG 254
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
255 CTGATGAGCAGACCATGAGAGGTGAGGCTTCAATCGAATCGGAGTGGAGAACACTG 314
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
315 ACCCGGTGGCGGAGAGACGGGGCAGCGCTGCTCAAGAGACTCAGAGCGCGCGAGGCC 374
OY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyLysVal 140
375 CGCGTGGCGCGGACATGAGAGACTGTCGGCCGCTGTCAGTACCGCGCGCAGAGTG 434
OY 141 GlnAlaMetLeuGlyGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
435 CAGGCGCATGCTCGCCAGAGCACCGAGAGCTGGCGGCTGCTCCACCTCGGCC 494
OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
495 AAGCTGCGTAAGCGGCTCTCCGCGATGCGCATGACTGTCAGAACCGCTGGCACTGTAC 554
OY 181 GlnAlaGlyAlaArgGluGlyAlaGlnArgGlyLeuSerAlaIleArgGluArgLeuGly 200
555 CAGCGCGGGGGCGCGGAGGGCGCCAGCGCGGCTCAGCGCATCCGGAGAGCGCTGGGG 614
OY 201 ProLeuVal 203
615 CCCCTGTGTG 623
Db

RESULT 3
AX302545
LOCUS AX302545 1147 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 63 from Patent WO01/5177.
ACCESSION AX302545
VERSION AX302545.1 GI:17383082
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Morin, P.J., Sherman-Bauslt, C.A., Pizer, E.S. and Hough, C.D.
TITLE Tumor markers in ovarian cancer
JOURNALS Patent: WO 01/5177-A 63 11-OCT-2001.
THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)
FEATURES
Location/Qualifiers
1..1147
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 210 a 365 c 425 g 147 t
ORIGIN

Alignment Scores:
Pred. No.: 1.59e-71 Length: 1147
Score: 1014.00 Matches: 203
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x AX302545 (1-1147)

OY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
46 ATGAAGGTTCTGTGGGCTGCGCTGTGGTGCATCTCCGAGATCCAGATGCCAGGCCAAGGTG 105
OY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
106 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGGCCAGACAGACGAGTGGAGAGC 165
OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
166 GGGCAGCCGCTGGGAACCTGGCACTGGCTCTTTGGGATTACCTCGCTGGGTGCAGACA 225
OY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
226 CTGCTGAGCAGAGTCAGAGAGAGCTGCTCAGCTCCAGAGTCACCCAGGAACCTGAGGGCG 285
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
286 CTGATGAGCAGACCATGAGAGGTGAGGCTTCAATCGAATCGGAGTGGAGAACACTG 345
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
346 ACCCGGTGGCGGAGAGACGGGGCAGCGCTGCTCAAGAGACTCAGAGCGCGCGAGGCC 405
OY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyLysVal 140
406 CGCGTGGCGCGGACATGAGAGACTGTCGGCCGCTGTCAGTACCGCGCGCAGAGTG 465
OY 141 GlnAlaMetLeuGlyGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
466 CAGGCGCATGCTCGCCAGAGCACCGAGAGCTGGCGGCTGCTCCACCTCGGCC 525
OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
526 AAGCTGCGTAAGCGGCTCTCCGCGATGCGCATGACTGTCAGAACCGCTGGCACTGTAC 585
OY 181 GlnAlaGlyAlaArgGluGlyAlaGlnArgGlyLeuSerAlaIleArgGluArgLeuGly 200
586 CAGCGCGGGGGCGCGGAGGGCGCCAGCGCGGCTCAGCGCATCCGGAGAGCGCTGGGG 645
OY 201 ProLeuVal 203
646 CCCCTGTGTG 654
Db

RESULT 4
BD004278
LOCUS BD004278 1156 bp DNA linear PAT 31-JAN-2002
DEFINITION Apo E humanized mammal.
ACCESSION BD004278
VERSION BD004278.1 GI:18632239
KEYWORDS JP 2001017028-A/2.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Fujita, S., Hamanaka, H., Fukui, Y. and Yokoyama, M.
TITLE Apo E humanized mammal
JOURNALS Patent: JP 2001017028-A 2 23-JAN-2001;
MITSUBISHI CHEMICAL CORP

COMMENT OS Homo sapiens (human)
PN JP 2001017028-A/2
PD 23-JAN-2001

PE 28-APR-2000 JP 2000128919
PR
PI SHINOBU FUJITA, HIROKI HAMANAKA, YOKO FUKUI, MINESUKE YOKOYAMA PC
A01K67/02T, A61K45/00, A61P25/28, A61P43/00, C12N5/10, PC
C12N5/09//C07K14/775
PC (C12N5/10, C12R1:91), C12N5/00, C12N15/00, (C12N5/00, C12R1:91) CC

FEATURES	source	FT	Key	CDs	Location/Qualifiers
BASE COUNT			208	a	432 g 148 t
ORIGIN					

QY	141	GINA1aM6tLEuGLyGLSerThrcInGLuLeuAaVgVA1aArgLEuA1SerIstLeuAry	160
Db	481	CAGGCATGCTCGGCCAAGCACCGAGAGAGCTCGCGGGTCCGCTCCACCTGCGC	540
QY	161	LysLeuAaArgLySAaArgLEuAaArgSApA1aAsPasPLeuGLnLySAaArgLEuA1aValYr	180
Db	541	AAGCTGCGTAAGCGGCTCCTCCGCGATGCGCATGACCTGCAGAAACGCGCTGGCACTGAC	600
QY	181	GINA1aGLyA1aAaArgGLnGLyA1aGLuAaArgLyLeuAaArgLEuA1aArgLEuAaArgLEuGLy	200
Db	601	CAGCGCGGGGCCCCGAGAGGCGCCAGAGCGGCGCTCAGCGCCATCCGCGAGCGCTGGG	660
QY	201	ProLeuVal	203
Db	661	CCCCGTGGT	669
RESULT 6			
LOCUS	BC003557	1186 bp	mRNA linear PRI 12-JUL-2001
DEFINITION	Homo sapiens, apolipoprotein E, clone MGC:1571 IMAGE:3355712, mRNA, complete cds.		
ACCESSION	BC003557		
VERSION	BC003557.1	GI:13097698	
KEYWORDS	MGC.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 1186) Strausberg, R. Direct Submision Submitted (20-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
AUTHORS	NIH-MGC Project URL: http://mgc.ncl.nih.gov Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@cgsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabh, Parvaneh Seedi, Jacqueline Schein, Duane Smalius, Michael Smith, Lorraine Spence, Jeff Stolt, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.		
REMARK	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: http://image.llnl.gov Series: IRAL Plate: 6 Row: h Column: 23 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4557324.		
COMMENT	Location/Qualifiers 1..1186 /organism="Homo sapiens" /db_xref="LocusID:348" /db_xref="taxon:9606" /clone="MGC:1571 IMAGE:3355712" /tissue_type="Eye, retinoblastoma" /clone_11p="N1H_MGC_16" /lab_host="DH10B-R" /note="Vector: pOTB7" 46..999 /codon_start=1 /product="apolipoprotein E" /protein_id="AAH03557.1"		
FEATURES	SOURCE		

[illegible]

REFERENCE
AUTHORS
TITLE
JOURNAL

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;
Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo.
1 (pases 1 to 1110)
Mimoto, H. and Teranishi, Y.
METHOD FOR PRODUCING APOLOPROTEIN
Patent: JP 1994315392-A 1 15-NOV-1994;

COMMENT	OS	Homo sapiens (human)
PN	JP	1994315392-A/1

FT	source	1.110	/organism='Homo sapiens'
FT		15.68	
FT	sig_peptide	69.968	/product='human apolipoprotein E3'
FT	mat_peptide		

Alignment Scores:	
Pred. No.:	5,47e-71
Score:	1007.00
Percent Similarity:	99.51%
Best Local Similarity:	99.51%
Query Match:	99.31%
DB:	6
Gaps:	0
Length:	1110
Matches:	202
Conservative:	0
Mismatches:	1
Indels:	0
Gaps:	0

US-09-827-854-15_COPY_1-203 x E08423 (1-1110)

OY	1	MetysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaVal	20
Db	15	ATGAAGTTCTGTGGCGCTGCTTCTGGTCACTTCTGCGAGATCCAGCCAAAGTG	74
OY	21	GIuGlnAlaValGIuThrGIuProGIuProGIuLeuArgGlnGlnThrGIuTrpGlnSer	40
Db	75	GAGCAAGCGGTGGAGACAGACCAGGACCCGAGCTGGCCAGCAGACCGAATGGCAGAGC	134
OY	41	GIyGlnArgTrpGIuLeuAlaLeuGIuArgPheTrpAspIlyLeuArgTrpValGlnThr	60
Db	135	GGCCAGCGCTGGGAACGTGGCACTGTGGATTAACCTGCCTGGGTGCACACA	194
OY	61	LeuSerGIuGlnValGlnGIuLeuLeuSerSerGlnValThrGlnGIuLeuArgAla	80
Db	195	CTGTCTAGCAGAGTGCAGAGAGAGACTCTCACTCCAGGTACACCAGCAACTAGAGGCG	255
OY	81	LeuMetAspGIuThrMetIysGIuLeuIuAlaTrpIysSerGIuLeuGIuGlnLeu	100
Db	255	CTGATGGACGAGACCATTGAAGAGATTAAAGCCCTTCAAAATCTGCAACTGGAGGAACAACCTG	314
OY	101	ThrProValAlaGIuGIuThrArgAlaArgLeuSerIysGIuLeuGlnAlaAlaGlnAla	120
Db	315	ACCCCGGTGGCGGAGAGACCGGGGCAAGGCTGTCCAAAGGAGGTGCAGAGCGGGCGCAAGCC	374
OY	121	ArgLeuGlnAlaAspMetGIuAspValCysGIuArgLeuValGlnTrpArgGIuGlnVal	140
Db	375	CGCGTGGGCGCGAGCATGGAGAGAGCTGGCGGGCGCTGTGGTCACTACCGGGGCGAGGTG	433
OY	141	GlnAlaMetLeuGIuGlnSerThrGIuLeuArgValArgLeuAlaSerHisLeuArg	160
Db	435	CAGGCCATGGTGGCCAGAGACCGAGAGAGCTGGGGGTGGCGCTCCCTCCACACTGGC	494

Oy	161	LysLeuAArgLysAArgGLeuLeuAArgAspAlaAspSpleuGlnLysAArgLeuAlaValTyr	180
Db	495	AAGCGCGTAAGCGGGCTCTCCGGGATGCGCATACCTGCAGAAAGCCCTGGCAGTGTAC	554
Oy	181	GlnAlaGlyAlaAArgGluGlyAlaGlnAArgGlyLeuSerAlaIleAArgGluAArgGly	200
Db	555	CAGGCCGGGGGCCCGGAGGGGCCCGAGGCGGGCTTCAAGCGCCATCCCGGAGCGGCTGGGG	614
Oy	201	ProLeuVal	203
Db	615	CCCTCGGTG	623

LOCUS	BD004277	1156 bp	DNA	linear	PAT 31-JAN-2002
DEFINITION	Apo E humanized mammal.				
ACCESSION	BD004277				
VERSION	BD004277.1	GI:18632238			
KEYWORDS	JP 2001017028-A/1.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				

REFERENCE
AUTHORS
TITLE
JOURNAL

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1156)
Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.
Apo E humanized mammal
Patent: JP 2001017028-A 1 23-JAN-2001;

COMMENT	OS	Homo sapiens (human)

PI SHINOBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA PO
A0167/027, A61K45/00, A61P25/28, A61P43/00, C12N5/10, PC
C12N15/09, /C07K14/775,
PC (C12N5/10, C12R1: 91), C12N5/00, C12N15/00, (C12N5/00, C12R1: 91) CC

FH	Key	Location/Qualifiers
FT	CDS	
		(61). .(1011).

BASE COUNT	208 a	367 c	432 g	149 t
ORIGIN				

Alignment Scores:	
Pred. No.:	6,866-71
Score:	1006.00
Percent Similarity:	99.51%
Best Local Similarity:	99.51%
Query Match:	99.21%
	Length: 1156
	Matches: 202
	Conservative: 0
	Mismatches: 1
	Indels: 0

US-09-827-854-15_COPY_1_203 (1-203) x BD004277 (1-1156)

OY	1	MelystValLeuThrPAlaIaIaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIaIaVal	20
Db	61	ATGAAGGCTTCGTGTGGGCTGGCTCTGGTCAATTCCTGGAGGATCCAGGCCAAGGTG	120
OY	21	GIuGlnAlaValAlGIuThrGluProGIuProGIuLeuArgGlnGlnThrGIuTPGIuSer	40
Db	121	GAGCAAGGGGTGGAGACACAGACCAGCCGAGCTGGCCACGACACCCAGATGGCAGACC	180
OY	41	GlyGlnArgTTPGIuLeuAlaIaLeuGlyArgPheThrPaspTyrLeuArgTTPValGlnThr	60
Db	181	GGCCAGCGCTGGGAACCTGGCACTGGGTCGCTTTTGGGATTAACCTGGCGCTGGGTGCAGACA	240
OY	61	LeuSerGIuGlnValGlnGluGluLeuLeuSerGlnValThrGlnGluLeuArgAla	80
Db	241	CTGTCTGACAGGTGCAGGAGAGACTCTCAGCTCCAGAGTGCACCCAGGAACATAGGGCG	300


```

DB 122 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCGCCAGACAGACGAGTGGCAGAGC 181
QY 41 GYGLNARTRGTLNLEUALALEUGLYARPHETTPASPTYLLEUARTRPVALGLNTHR 60
DB 182 GGCACAGCGCTGGAACTGGCACTGGTCCCTTTTGGATTTACTCTCGCTGGGTGACAGCA 241
QY 61 LEUSERGLNGLNVALGLNGLNGLNLEUUSERSERGLNVALTHRGlnLEUENURGALA 80
DB 242 CTGTCTGAGCAGGTGACAGAGAGCTGCTCAGCTCCCAAGTACCCAGAACTGAGAGCG 301
QY 81 LEUMETASPGLTUThMETYSGLULEUYSALAITYLSERSERGLNLEUGLNUGLNLEU 100
DB 302 CTGATGAGACAGAACCATGAGAGAGTTGAAGCGCTTCAAAATCGGAATCGAGAGAACACTG 361
QY 101 THRPROVALAAGLUGLNUThRARGLAARGLEUSERLYSGLUENGLNLAALAGLNA 120
DB 362 ACCCGGTAGCGGAGAGAGCGCGGACGCTGTCCAAAGACTGACAGCGGCGAGGCG 421
QY 121 ARGLEUGLYALAASPMETGLUASPVALCYSGLYATRGLEUVALGLNTRYARGLYGLNU 140
DB 422 CGGCTGGGCGGACATGAGAGAGCTGTCGCGCCGCTGTGCTACAGTACCGCGGCGAGTG 481
QY 141 G1NALAMETLEUGLYGlnSERThRGlnLEUENARGLAARGLEUALASERHISLEUAR 160
DB 482 CAGGCGCATGCTCGCGCAGACACCGAGAGCTGCGGCTGCGCTCCACCTGCGC 541
QY 161 LYSLEUARGLYSARGLLEUENURGASPALAASPAPLEUGLNUARGLEUALAVALTYR 180
DB 542 AAGCTGCGTAACGGCTCTCCGCGATCCCGATGACCTGCAGAACCGCTGCGACTGTAC 601
QY 181 G1NALAGLYALAARGGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLN 200
DB 602 CAGGCGCGGCGCGCGGAGGCGCGCCAGCGCGGCTCAGCGCATCCGCGAGCGCTCGG 661
QY 201 PROLEUVAL 203
DB 662 CCCCTGGTG 670

```

RESULT 11

```

LOCUS 115975 1157 bp DNA linear PAT 03-APR-1996
DEFINITION Sequence 1 from patent US 5472858.
ACCESSION 115975
VERSION 115975.1 GI:1250883
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1157)
AUTHORS Attie,A.D., Gretsch,D.G., Shurley,S.L. and Beckage,N.E.
TITLE Production of recombinant proteins in insect larvae
JOURNAL Patent: US 5472858-A 1 05-DEC-1995;
FEATURES
source
1..1157
/organism="unknown"

```

```

BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN

```

Alignment Scores:

```

Pred. No.: 8,24e-71 Length: 1157
Score: 1005.00 Matches: 201
Percent Similarity: 99.01% Conservative: 0
Best Local Similarity: 99.01% Mismatches: 2
Query Match: 99.11% Indels: 0
DB: 6 Gaps: 0

```

```

US-09-827-854-15_COPY_1_203 (1-203) x 115975 (1-1157)

```

```

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 62 ATGAAGGTTTGTGGTGGCTGCTGTCATTCCTGGCGAGATGCGAGGCAAGGTG 121

```

```

QY 21 G1UGLNALAVALGLNThRGlnUPROGLNLEUENURGALGlnThRGlnUPRGLNser 40
DB 122 GAGCAAGCGGTGGAGACAGAGCCGAGCGAGCTGCGCCAGACAGACGAGTGGCAGAGC 181
QY 41 GYGLNARTRGTLNLEUALALEUGLYARPHETTPASPTYLLEUARTRPVALGLNTHR 60
DB 182 GGCACAGCGCTGGAACTGGCACTGGTCCCTTTTGGATTTACTCTCGCTGGGTGACAGCA 241
QY 61 LEUSERGLNGLNVALGLNGLNGLNLEUUSERSERGLNVALTHRGlnLEUENURGALA 80
DB 242 CTGTCTGAGCAGGTGACAGAGAGCTGCTCAGCTCCCAAGTACCCAGAACTGAGAGCG 301
QY 81 LEUMETASPGLTUThMETYSGLULEUYSALAITYLSERSERGLNLEUGLNUGLNLEU 100
DB 302 CTGATGAGACAGAACCATGAGAGAGTTGAAGCGCTTCAAAATCGGAATCGAGAGAACACTG 361
QY 101 THRPROVALAAGLUGLNUThRARGLAARGLEUSERLYSGLUENGLNLAALAGLNA 120
DB 362 ACCCGGTAGCGGAGAGAGCGGCGACGCGCTGTCCAAAGACTGACAGCGGCGAGGCG 421
QY 121 ARGLEUGLYALAASPMETGLUASPVALCYSGLYATRGLEUVALGLNTRYARGLYGLNU 140
DB 422 CGGCTGGGCGGACATGAGAGAGCTGTCGCGCCGCTGTGCTACAGTACCGCGGCGAGTG 481
QY 141 G1NALAMETLEUGLYGlnSERThRGlnLEUENARGLAARGLEUALASERHISLEUAR 160
DB 482 CAGGCGCATGCTCGCGCAGACACCGAGAGCTGCGGCTGCGCTCCACCTGCGC 541
QY 161 LYSLEUARGLYSARGLLEUENURGASPALAASPAPLEUGLNUARGLEUALAVALTYR 180
DB 542 AAGCTGCGTAACGGCTCTCCGCGATCCCGATGACCTGCAGAACCGCTGCGACTGTAC 601
QY 181 G1NALAGLYALAARGGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLN 200
DB 602 CAGGCGCGGCGCGCGGAGGCGCGCCAGCGCGGCTCAGCGCATCCGCGAGCGCTCGG 661
QY 201 PROLEUVAL 203
DB 662 CCCCTGGTG 670

```

RESULT 12

```

LOCUS HUMAPOE 1157 bp mRNA linear PRI 08-AUG-1995
DEFINITION Human apolipoprotein E mRNA, complete cds.
ACCESSION M12529
VERSION M12529.1 GI:178848
KEYWORDS
SOURCE Homo sapiens (clone: PHAE[112,178,813].) male 57-year old liver
CDNA to mRNA.
ORGANISM Homo sapiens

```

```

REFERENCE 1 (bases 1 to 1157)
AUTHORS Mammalla; Euteria; Primates; Catarrhini; Homnidae; Homo.
McLean,J.W., Elshoubagy,N.A., Chang,D.J., Mahley,R.W. and
Taylor,J.M.

```

```

TITLE Human apolipoprotein E mRNA. CDNA cloning and nucleotide sequencing
of a new variant
JOURNAL J. Biol. Chem. 259 (10), 6498-6504 (1984)
MEDLINE 84212473
PUBMED 6327682

```

```

FEATURES
source
1..1157
Location/Qualifiers

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19q13.2"
/clone="PHAE[112,178,813]."
/sex="male"
/tissue="liver"
/dev_stage="57-year old"
1..1157
/gene="APOE"
<1..1157
/gene="APOE"

```



```

CDS
/note="apoE mRNA: G00-119-691"
62..1015
/gene="APOE"
/note="precursor"
/codon_start=1
/product="apolipoprotein E"
/protein_id="AAB59518.1"
/db_xref="GI:178849"
/db_xref="GDB:G00-119-691"
/translation="MKVLAALLVPLPGLCAQKAEVAVETPEPELRQOTEMOGGWR
ELALGFMVLRWVQTLSEVOEELISVQTOELRLMETKELKAYSPLEOLTP
VAETRLARSKELQTAQARLQADMEVCGRLVOYRGEVQAMIGOSTEEELRLASLR
KLRRLRDPDDIQRLAVYQAGARGARGLSALRELRGLVEGGRVAAVGSILAG
OPLOERAAQMGSERLRAMDEKSGRTDRLEDEKQVAVRKLIEBOAOQIRLOAEAFQ
ARKSWFEPLVEDMOROMAGLVEKVAAGVTSAAVPSPDNH"
62..115
/gene="APOE"
/note="APOE"
/note="G00-119-691"
116..1012
/gene="APOE"
/product="apolipoprotein E"
/note="G00-119-691"
370
/gene="APOE"
/note="a in PHAE[112,178]; g in PHAE813; G00-119-691"
/replace="g"
410
/gene="APOE"
/note="a in PHAE[112,178]; g in PHAE813; G00-119-691"
/replace="g"
569
/gene="APOE"
/note="c in PHAE[112,178]; g in PHAE813; G00-119-691"
/replace="g"
784
/gene="APOE"
/note="c in PHAE[112,178]; c in PHAE813; G00-119-691"
/replace="c"
859
/gene="APOE"
/note="g in PHAE[112,178]; a in PHAE813; G00-119-691"
/replace="a"
BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN 30 bp upstream of BstNI site.

Alignment Scores:
Pred. No.: 8,24e-71 length: 1157
Score: 1005.00 Matches: 201
Percent Similarity: 99.01% Conservative: 0
Best Local Similarity: 99.01% Mismatches: 2
Query Match: 99.11% Indels: 0
DB: 9 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x HUMAPOE (1-1157)
QY 1 MetlysValleuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
DB 62 ATGAAGGTCTGTGGGTGCTGTGCATCTCTGCGCAGAGATGCCAGCGCAAGTGT 121
|||||
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrGlnSer 40
|||||
DB 122 GAGCAAGCGGTGAGACAGAGCGGAGCCGAGCTGCGCAGACAGCCGAGTGGCAGAGC 181
|||||
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
DB 182 GGGCAGCGCTGGGAACTGGCGACTGGCTTTGGGATTACTCGCGCTGGTGGCAGACA 241
|||||
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
DB 242 CTGCTGAGCAGGTGCGAGGAGCTGCTCAGCTCCCAAGTACCAAGAACTGAGGGCG 301
|||||
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
|||||

```

```

DB 302 CTGATGACGACGACCATGAAGAGTTGAMAGCGCTACAAATCGGAACTGGAGAACACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
|||||
DB 362 ACCCGGTACCGAGGAGAGACGGCGGCGCTCTCCAGAGAGTGGACAGCGGCGAGGCC 421
|||||
QY 121 ArgLeuGlnAlaAspMetGluAspValGlyGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
DB 422 CGGCTGGCGCGGACATGAGGAGCTGTGGCGCGCGCTGTGCAGTACCGCGCGGAGTG 481
|||||
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
DB 482 CAGCGCATGCTCGGCCACAGACCGAGAGAGCTCGGGTGGCTCCGCTCCACCTGGCG 541
|||||
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
DB 542 AACCTGCGTAGAGGGCTCTCCGCGCATCCGATCGACCTGAGAAAGCGCTGGCAGTGC 601
|||||
QY 181 GlnAlaGlyAlaArgGluGluGluArgGlyLysSerAlaIleArgGluArgLeuGly 200
|||||
DB 602 CAGCGCGGCGCGCGCAGAGCGCGCAGCGCGCTCAGCGCATCCGCGAGCGCTGGGG 661
|||||
QY 201 ProLeuVal 203
DB 662 CCCCTGGTG 670

RESULT 13
LOCUS BD004279 1156 bp DNA linear PAT 31-JAN-2002
DEFINITION Apo E humanized mammal.
ACCESSION BD004279
VERSION BD004279.1 GI:18632240
KEYWORDS JP 2001017028-A/3.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1156)
AUTHORS Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.
TITLE Apo E humanized mammal
JOURNAL Patent: JP 2001017028-A 3 23-JAN-2001;
MITSUBISHI CHEMICAL CORP
OS Homo sapiens (human)
PN JP 2001017028-A/3
PD 23-JAN-2001
PF 28-APR-2000 JP 2000128919
PR SHINOBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA PC
A01K67/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10, PC
C12N15/09//C07K14/775,
PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,(C12N5/00,C12R1:91) CC

FEATURES
source FT CDS Location/Qualifiers
1..1156
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 208 a 369 c 432 g 147 t
ORIGIN

Alignment Scores:
Pred. No.: 1,42e-70 length: 1156
Score: 1002.00 Matches: 202
Percent Similarity: 99.51% Conservative: 0
Best Local Similarity: 99.51% Mismatches: 1
Query Match: 98.82% Indels: 0
DB: 6 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x BD004279 (1-1156)
QY 1 MetlysValleuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||

```


[illegible]

ALIGNMENT SCORES:	ALIGNMENT SCORE:	PERCENT SIMILARITY:	BEST LOCAL SIMILARITY:	QUERY MATCH:	DB:
2e-69	984.00	98.99%	98.99%	97.04%	6
Length: 660	Matches: 197	Conservative: 0	Mismatches: 2	Indels: 0	Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x A62340 (1-660)

QY	1	MELTVSVALETTTPALALALEUENVALTHRPHLEUALAGLYCYSGLNALALYSVAL	20
DB	62	ATGAAGGTTCTTGTGGGCTGCGTTGCTGGTCACATTCCTGGCAGATGCGCAGGCCAAGGCG	121
QY	21	GLUGINALAVALAIGLTHRCINUPROGILUPROGILUENUARGINLGINTHRCINUPGINSER	40
DB	122	GAGCAAGCGGTGGAGACAGAGGCCGAGCCGAGCTGCGCCAGCAACCGAGTGGCAGAGC	181
QY	41	GLIGINARITRPGILUENUALALEUENLYARQPHETIPASPTRYLEUARGITRYVALGINTHR	60
DB	182	GGCCAGCGCTGGGAAGTGGCACTGGCACTGGGTGGCTTTGGATTTACCTGCGCGGTGCAGACA	241
QY	61	LEUSERGLUGINVALGINSGLUENLEUSERSERGINVALTHGLINGLUENUARGALA	80
DB	242	CTGTCTGAGACAGTGCAGAGAGAGCTGCTCAGCTCCCAAGTCACCCAGAACTGAGGGCG	301
QY	81	LEUMETASPGILUTHMETLYSGILUENULYSALATRYLYSSERGLINGLUENGLINGLINTLEU	100
DB	302	CTGATGGAGACAGACATGAGAGAGTTGAAGGCTTCACAAATCGGAATCGAGAGAACACTG	361
QY	101	THRPVALALAGLIGLUTHTRARGALALARGLEUSERLYSGILUENGINALALAGLINALA	120
DB	362	ACCCCGGTRAGCGAGAGACAGCGGGCAGCGCTGCTCCAGAGACTGCAGAGCGCGCAGGCC	421
QY	121	ARGLEUGLYALASPMEGLINUSPYALCYSGIYARGLEVALGINTRYARGIYGLIYVAL	140
DB	422	CGCGTGGGGCGGACATGGAGAGACGTGGCGGCGCCCTGCTGTCACACTACCGCGGCGAGGTG	481
QY	141	GINALAMETLEUGLYGINSERTHRGLINGLUENUARGVALARGLEUALASERTHISLEUARG	160
DB	482	CAGGCCATGCTCGCGCAGAGCACCCAGAGAGCTGCGGCTGCGCTCCACCTCGCGC	541
QY	161	LYSLIEATRGLYSARGLEUENUARGSPALIASPASPLEUGINLYSARGLEUALALAYLR	180
DB	542	AAGCTCGGTAAGCGGGCTCTCCGCAATCCCGATGACCTGCAGAAACGGCTGGCACTGTATC	601
QY	181	GINALAGLYALARGGLUGLYVALAGLUNARGGLYLEUSERALALALARGLUNARGLEU	199
DB	602	CAGGCGGGGGCGCGGAGAGGCGCGCAGAGGCGGCTTCACAGCGCCATTCGCGGAGCGCGTG	658

RESULT 15	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
A62342	A62342	Sequence 30 from Patent WO9712992.	660 bp	DNA	linear	PAT 09-MAR-1998	
A62342	A62342						
A62342.1	GI:3716295						
		unidentified.					
		unidentified.					
		unclassified.					
		1 (bases 1 to 660)					
		Van, L.F., Burbach, J.P. and Grosvel, F.G.					
		DIAGNOSIS METHOD AND REAGENTS					
		Patent: WO 9712992-A 30 10-APR-1997;					
		ROYAL NETHERLANDS ACADEMY OF A (NL)					
		Other publication AU 7142796 970428.					
		Location/Qualifiers					
		1..660					
		/organism="unidentified"					
		/db_xref="taxon:32644"					

BASE COUNT 129 a 197 c 247 g 87 t

ORIGIN

Alignment Scores:

Pred. No.:	2e-69	Length:	660
Score:	984.00	Matches:	197
Percent Similarity:	98.99%	Conservative:	0
Best Local Similarity:	98.99%	Mismatches:	2
Query Match:	97.04%	Indels:	0
DB:	6	Gaps:	0

US-09-827-854-15_COPY_1_203 (1-203) x A62342 (1-660)

```
QY      1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
          |||
DB      62 ATGAAGTTCTGTGGGTGCTGCTGCTGCATTCCTCGCAGGATGCCAGGCCAAGGTG 121

QY      21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
          |||
DB      122 GACCAAGCGGTGAGAGAGCCGAGACCCGAGCTGCGCCAGCAGACCGAGTGGCAGAGC 181

QY      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
          |||
DB      182 GGCAGAGCGTGGGAGACGCGACCTGGCTTTGGGATTACCTCGCGCTGGTGCGAGACA 241

QY      61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
          |||
DB      242 CTGCTGAGCAGGTCGAGGAGAGCTGCTCAGCTCCCAAGTCAACCAGAACTGAGGGCG 301

QY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
          |||
DB      302 CTGATGACAGACCATGAGAGAGTTGAAAGCCCTACAAATCGGAAGTGGAGAACACTG 361

QY      101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
          |||
DB      362 ACCCGGTAGCGAGGAGACGCGGGCAGCGCTGTCCAAAGAGCTGCAGAGCGCGCAGGCC 421

QY      121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
          |||
DB      422 CGCCTGGGCGCGGACATGAGGAGCGTGTGCGCGCGCTGGTGCAGTACCAGGCGAGGTG 481

QY      141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
          |||
DB      482 CAGGCCATGCTCGGCCGAGACACGAGAGAGCTCGGGGTGCGCTCGCTCCACCTGCGC 541

QY      161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
          |||
DB      542 AAGCTGGTAAAGGGGCTCCCGCGATCCCGATCCGAGACCTGCAGAAAGCGCTGGCAGTGTAC 601

QY      181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeu 199
          |||
DB      602 CAGGCGGGGCGCGCGAGGCGCGGAGCGGCGCTCAGCGCATCCGCGAGCGCGCTG 658
```

Search completed: March 14, 2003, 17:32:22
Job time : 1481.5 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame-plus.p2n model

Run on: March 14, 2003, 12:08:17 ; Search time 114.454 Seconds
(without alignments)
3994.237 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_203
Perfect score: 1014
Sequence: 1 MKVLMALVTFIAGCOAKV.....AREGAEGRGLSAIRERLGPLV 203

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlp
-O=/cgn2.1/USPRO/spool/US09827854/runat_11032003_101609_27476/app-query.fasta.1.3576
-DB=N_genseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCCALIGN=200 -THR_SCORE=0 -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pro -NORR=ext -HEAPSIZE=500 -MILEN=0 -MAXLEN=200000000
-USER=US09827854.cgn.1.1.101 @runat_11032003_101609_27476 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
N_genseq_101002: *
1: /SID22/gcgdata/genseq/genseqn-emb1/NA1980.DAT: *
2: /SID22/gcgdata/genseq/genseqn-emb1/NA1981.DAT: *
3: /SID22/gcgdata/genseq/genseqn-emb1/NA1983.DAT: *
4: /SID22/gcgdata/genseq/genseqn-emb1/NA1983.DAT: *
5: /SID22/gcgdata/genseq/genseqn-emb1/NA1984.DAT: *
6: /SID22/gcgdata/genseq/genseqn-emb1/NA1985.DAT: *
7: /SID22/gcgdata/genseq/genseqn-emb1/NA1986.DAT: *
8: /SID22/gcgdata/genseq/genseqn-emb1/NA1987.DAT: *
9: /SID22/gcgdata/genseq/genseqn-emb1/NA1988.DAT: *
10: /SID22/gcgdata/genseq/genseqn-emb1/NA1989.DAT: *
11: /SID22/gcgdata/genseq/genseqn-emb1/NA1990.DAT: *
12: /SID22/gcgdata/genseq/genseqn-emb1/NA1991.DAT: *
13: /SID22/gcgdata/genseq/genseqn-emb1/NA1992.DAT: *
14: /SID22/gcgdata/genseq/genseqn-emb1/NA1993.DAT: *
15: /SID22/gcgdata/genseq/genseqn-emb1/NA1994.DAT: *
16: /SID22/gcgdata/genseq/genseqn-emb1/NA1995.DAT: *
17: /SID22/gcgdata/genseq/genseqn-emb1/NA1996.DAT: *
18: /SID22/gcgdata/genseq/genseqn-emb1/NA1997.DAT: *
19: /SID22/gcgdata/genseq/genseqn-emb1/NA1998.DAT: *
20: /SID22/gcgdata/genseq/genseqn-emb1/NA1999.DAT: *
21: /SID22/gcgdata/genseq/genseqn-emb1/NA2000.DAT: *
22: /SID22/gcgdata/genseq/genseqn-emb1/NA2001A.DAT: *
23: /SID22/gcgdata/genseq/genseqn-emb1/NA2001B.DAT: *
24: /SID22/gcgdata/genseq/genseqn-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1014	100.0	954	24	AAD26035	Human apolipoprotein E
2	1014	100.0	1110	7	AAN60409	Human apolipoprotein E
3	1014	100.0	1147	22	ABA83113	Human apolipoprotein E
4	1014	100.0	1156	22	AAF84315	Human apolipoprotein E
5	1014	100.0	1156	24	AAD22048	Human apolipoprotein E
6	1014	100.0	1279	22	AAS22437	Human apolipoprotein E
7	1010	99.6	1156	24	AAD22052	Human apolipoprotein E
8	1006	99.2	1156	22	AAF84314	Human apolipoprotein E
9	1006	99.2	1156	24	AAD22049	Human apolipoprotein E
10	1006	99.2	1156	24	AAD22051	Human apolipoprotein E
11	1005	99.1	1157	17	AAU06957	Human apolipoprotein E
12	1005	99.1	1157	24	ABN95746	Human apolipoprotein E
13	1005	99.1	1157	24	ABK64514	Human apolipoprotein E
14	1005	99.1	1157	24	ABL65450	Human apolipoprotein E
15	1004	99.0	1110	6	AAN50450	Sequence encoding
16	1002	98.8	1156	22	AAF84316	Human apolipoprotein E
17	1002	98.8	1156	24	AAD22047	Human apolipoprotein E
18	999	98.5	1156	24	AAD22050	Human apolipoprotein E
19	984	97.0	1156	18	AAF69792	Partial human apolipoprotein E
20	891.5	87.9	1107	19	AAU75756	Human apolipoprotein E
21	825.5	81.4	9360	24	AB131915	Human apolipoprotein E
22	825.5	81.4	10716	24	AAD26034	Human apolipoprotein E
23	820	80.9	3805	20	AAZ09524	Human apolipoprotein E
24	820	80.9	3805	20	AAZ09526	Human apolipoprotein E
25	803.5	79.2	10716	24	AAD26108	Human apolipoprotein E
26	760.5	75.0	965	24	AAD32081	Bovine Apoe gene.
27	760.5	75.0	5617	24	AAD32077	Human alpha-1-anti
28	760.5	75.0	6026	24	AAD32075	Human albumin prom
29	756	74.6	1381	22	AAS22673	Human apolipoprotein E
30	724.5	71.4	1126	19	AAV29159	Nucleotide sequenc
31	678	66.9	600	20	AAU98955	EST clone cp147.
32	651	64.2	407	24	ABK34238	Human apolipoprotein E
33	625	61.6	478	24	ABN95924	Human apolipoprotein E
34	625	61.6	478	24	ABL62679	Human apolipoprotein E
35	625	61.6	478	24	ABL67340	Human apolipoprotein E
36	625	61.6	478	24	ABL67340	Human apolipoprotein E
37	559	55.1	936	17	AAU18070	Human apolipoprotein E
38	559	55.1	936	17	AAU18070	Human apolipoprotein E
39	521	51.4	597	12	AAQ11980	Human apolipoprotein E
40	514	50.7	330	12	AAQ11980	Human apolipoprotein E
41	512	50.5	597	15	AAU69099	Human apolipoprotein E
42	504	49.7	405	21	AAU69099	Human apolipoprotein E
43	485	47.8	345	22	AAU69099	Human apolipoprotein E
44	419	41.3	260	21	AAA40342	Human apolipoprotein E
45	410	40.4	382	21	AAA44191	Mouse secreted exp

ALIGNMENTS

RESULT 1
AAD26035
AAD26035 standard; CDNA: 954 BP.
26-MAR-2002 (first entry)

Human apolipoprotein E (APOE) CDNA.
Human; anti-lipemic; neuroprotective; nontoxic; genetic variant; APOE; apolipoprotein E; haplotyping; familial dysbetalipoproteinemia; therapy; genotyping; type III hyperlipoproteinemia; Alzheimer's disease; SNP; atherosclerosis; single nucleotide polymorphism; chromosome 19q13.2; ss.

OS Homo sapiens.
XX Location/Qualifiers
FH key

FT	CDS	1..954
FT	variation	/product= "human APOE protein"
FT	variation	/replace (13, C)
FT	variation	/tag= a
FT	variation	/standard_name= replace (31, G)
FT	variation	/tag= b
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	/replace (364, A)
FT	variation	/tag= c
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	/replace (388, C)
FT	variation	/tag= d
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	/replace (498, G)
FT	variation	/tag= e
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	/replace (526, T)
FT	variation	/tag= f
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	/replace (622, A)
FT	variation	/tag= g
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "
PN	WO200179234-A2.	
PD	25-OCT-2001.	
PD	16-APR-2001; 2001WO-US12303.	
PR	14-APR-2000; 2000US-197188P.	
PA	(GENA-) GENMAISSANCE PHARM INC.	
PI	Choi JY, Kijem SE, Koshy B, Lee HH;	
PI	WPI: 2002-075064/10.	
DR	P-PSDB: AAEL5158.	
XX	Genotyping human apolipoprotein gene of individual for determining	
XX	haplotype of individual, involves determining identity of nucleotide	
XX	pair at specific polymorphic sites for two copies of gene -	
XX	Claim 26; Fig 2; 78pp; English.	
XX	The patent discloses novel genetic variants of human apolipoprotein	
XX	E (APOE) gene. The invention also relates to compositions and methods	
XX	for haplotyping and/or genotyping the APOE gene. The haplotyping	
XX	methods of the invention are useful for improving the efficacy and	
XX	reliability of several steps in the discovery and development of	
XX	drugs for treating diseases associated with APOE activity, e.g.	
XX	familial dysbetalipoproteinemia, type III hyperlipoproteinemia,	
XX	atherosclerosis, and Alzheimer's disease. They are useful to validate	
XX	APOE as a candidate agent for treating a specific condition or disease	
XX	predicted to be associated with APOE activity and in the design of	
XX	clinical trials of candidate drugs for treating a specific condition	
XX	or disease predicted to be associated with APOE activity. Genotyping	
XX	or haplotyping methods are useful to screen for compounds targeting	
XX	APOE to treat a specific condition or disease associated with APOE	
XX	activity. The present sequence is a cDNA encoding human APOE protein.	
XX	APOE gene is located on chromosome 19q13.2.	
XX	Sequence 954 BP; 168 A; 290 C; 377 G; 119 T; 0 other;	
XX	Alignment Scores:	
XX	Pred. NO.:	1,18e-87
XX	Score:	1014..00
XX	Percent Similarity:	100.00%
XX	Best Local Similarity:	100.00%
XX	Query Match:	100.00%
XX	DB:	24
XX	DB:	Gaps: 0

US-09-827-854-15_copy1_1_203 (1-203) x AAD26035 (1-954)	
QY 1 MettlysvallLeutPrAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20	
Db 1 ATGAAGGTTCTGTGGGCTGGCTGTGTGTACATTCTCGGAGAGATCCAGGCAAGGTG 60	
QY 21 GtUGlnAlaValAlaGluThrGluProGluProLuleuAArgGlnGlnThrGluTrpGlnSer 40	
Db 61 GAGCAAGCGGCTGGAGACACAGAGCCGAGCCCAAGCTGCCGCCAGCAGAGCTGGCAAGC 120	
QY 41 GlyGlnArGTrpGluLeuAlaLeuGlyArGpHeTrpAspTrpLysArGTrpValGlnThr 60	
Db 121 GGCCAGCGCTGGAGAACTGGCACTGGGTGCTTTTGGATTACTGCGCTGGTGCAGACA 180	
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerGerglnValThrGlnGluLeuArAla 80	
Db 181 CTGTCTGAGCAGGTGCAAGAGAGAGCTGCTCAGTCCAGGTCACCCAGAACTGAGGGCG 240	
QY 81 LeuMeTAspGluThrMetLysGluLeuLysAlaTyrrLysSerGluLeuGluGlnLeu 100	
Db 241 CTGATGACGAGACCATGTAAAGAGTTGAAGCCTTACAAATCGAACTGGAGAAACAATG 300	
QY 101 ThrProValAlaGluGluThrArGAlaArGLeuSerLysGluLeuGlnAlaAlaGlnAla 120	
Db 301 ACCCGGTGGGGGAGAGAGACCGCGGCAACGCTGTCCAAGAGCTGCAGCGCGGCAAGCC 360	
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArGLeuValGlnTyrrArGlyGluVal 140	
Db 361 CGGCTGGGCGGCGACATGTGAAGAGCTGTGCGGCGCGCTGTGGCAGTACC CGCGGAGGTG 420	
QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArGValArgLeuAlaSerHisLeuArg 160	
Db 421 CAGGCATGCTCGGCGCAAGACACCGAGAGCTGCGGGTGCCTCGCTCCACCTGCAC 480	
QY 161 LysLeuArGlyArGLeuLeuArGAspAlaAspAspLeuGlnLysArGLeuAlaValTyr 180	
Db 481 AAGCTGCCTAAGCGGCTCTCTCCGGATCCCATGACTCGCAAGAGCGCTTCGCAGTCTAC 540	
QY 181 GlnAlaGlyAlaArGLeuGlyAlaGluArGlyLeuSerAlaLeuArgGluArGLeuGly 200	
Db 541 CAGCGCGGCGCGCGCGAGAGCGCGCGAGCGGCTCTCAGCGCATCCGCGAGCGCTGGG 600	
QY 201 ProteuVal 203	
Db 601 CCCCTGGTG 609	
RESULT 2	
AAN60409	
AAN60409 standard; DNA, 1110 BP.	
XX AAN60409;	
XX AC	
XX DT 01-JAN-1980 (first entry)	
XX DE Human apolipoprotein-E.	
XX KW Apolipoprotein-E; hyperlipidemia; arteriosclerosis; ss.	
XX OS Homo sapiens.	
XX FH Key	Location/Qualifiers
XX FT CDS	15..968
XX PN AUB547513-A.	/*tag= a
XX PD 24-APR-1986.	
XX PF 17-SEP-1985;	85AU-0047513.
XX PR 11-JUN-1985;	85JP-0126989.
XX PA (MITU) MITSUBISHI CHEM IND KK.	

QY	181	GlnAaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200
Db	601	CAGCGCGGGGCGCCGAGGAGCGCGGACCGCGGCGCTCAGGCCATCGCGAGGCGCTGGGG	660
OY	201	ProLeuVal 203	
Db	661	CCCCCTGGTG 669	
RESULT 5			
AAD22048			
ID	AAD22048	standard; DNA; 1156 BP.	
XX			
AC	AAD22048;		
XX			
DT	12-FEB-2002	(first entry)	
XX			
DE	Human apolipoprotein E (apoE) isoprotein, apoE3 DNA.		
XX			
KM	Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;		
XX	hypertriglyceridaemia; low density lipoprotein; LDL; ds.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	61..1014	
FT		/*tag= a	
FT		/product= "Human apoE isoprotein, apoE3"	
FT	sig_peptide	61..114	
FT		/*tag= b	
FT	mat_peptide	115..1011	
FT		/*tag= c	
FT		/product= "Mature human apoE isoprotein, apoE3"	
XX			
PN	WO200177136-A1.		
XX			
PD	18-OCT-2001.		
XX			
XX	06-APR-2001; 2001WO-US11358.		
PR	06-APR-2000; 2000US-0544386.		
PR	04-OCT-2000; 2000US-0679088.		
PR	05-APR-2001; 2001US-0827854.		
XX			
XX	(KOSP-) KOS PHARM INC.		
PA	(UYBO-) UNIV BOSTON.		
XX			
PI	Zannis VI, Kypros KE;		
XX			
DR	WPI; 2002-010885/01.		
DR	P-PSDB; AAE13294.		
XX			
PT	New apolipoprotein E polypeptide and nucleic acid, useful for lowering		
PT	cholesterol, delaying the onset of or treating atherosclerosis in		
PT	mammal, without inducing hypertriglyceridaemia		
XX			
PS	Claim 14; Page 81; 91pp; English.		
XX			
CC	The present sequence is a human apolipoprotein E (apoE)		
CC	isoprotein, apoE3 DNA. The apoE lipoproteins are useful for		
CC	lowering cholesterol, delaying the onset of atherosclerosis,		
CC	treating or regressing atherosclerosis without inducing		
CC	hypertriglyceridaemia, in a mammal lacking an endogenous,		
CC	normally functioning apoE gene for low density lipoprotein (LDL)		
CC	receptor or is at risk for developing atherosclerosis due to		
CC	accumulation of lipoprotein remnants in the bloodstream or having		
CC	a defect in remnant removal.		
XX			
SQ	Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other;		
Alignment Scores:			
Pred. NO.:	1,49e-87	Length:	1156
Score:	1014.00	Matches:	203

	Percent Similarity:	100.00%	Conservative:	0
	Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%		Indels:	0
DB:	24		Gaps:	0
US-09-827-854-15_COPY_1_203 (1-203) x AMD22048 (1-1156)				
QY 1 MetLysValLeuTrpPalaIaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIaIaVal 20				
Db 61 ATGAAGCTTCGTGTGGCGCTGCTTCTGTCATATTCCTGGCAGATGCACGACCAAGGTG 120				
QY 21 GluGlnAlaValaIgiUthrGlnUthrGlnUthrGlnUthrGlnUthrGlnUthrGlnUthr 40				
Db 121 GAGCAAGCGGTGGAGACAGACAGCCGAGCCCGACCTGGCCAGCAGACCAAGTGGCAGAC 180				
QY 41 GlyGlnAqgTrpGluLeuAlaLeuGlyAArgPheTrpAspTyrLeuArgTrpValGlnThr 60				
Db 181 GGCACGCGCTGGGAACAGCTGGCTGGCTTTTGGGATTAACCTGGCTGGGTGGCAGACA 240				
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80				
Db 241 CTGTGTGACGAGGTGACAGAGAGACCTGCTCAGCTCCAGGTCCACCGAAGACTGAGAGGCG 300				
QY 81 LeuMetAspGluUthrMetLysGlnLeuLysAlaTyrLysSerGlnUthrGlnGlnLeu 100				
Db 301 CTGATGGACGAGACCATGAAAGAGACTTGAAGCCCTACAAATCGGAACCTGAGAACCACTG 360				
QY 101 ThrProValaIaGlnGlnUthrArgAlaAArgLeuSerLysGlnUthrGlnAlaAlaGlnAla 120				
Db 361 ACCCGCGTGGCGGAGAGACACCGCGGACCGCTGCCAAGAGACTGACAGCGCGCAGCGC 420				
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyAArgLeuValGlnThrArgGlyGlnVal 140				
Db 421 CGGCTGGCGCGGACATGAGAGAGCTGTGGCGCGCCTGGTGCAGTACCGCGGCGAGGTG 480				
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160				
Db 481 CAGGCGACATGCTGGCGCCAGACACCGAGAGACTCGGGTGGCGCTCCGCCACCACTGGCGC 540				
QY 161 LysLeuAArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180				
Db 541 AAGCTGCGCTAAGCGCGCTCTCTCCGCGATGCCGATGACCTGCAGAGCGCTTGGCAGTGTAC 600				
QY 181 GlnAlaGlyAlaAArgGlnGlnGlnAlaArgGlyLeuSerAlaIleArgGlnUthrGlnGly 200				
Db 601 CAGGCGGGGCGCGCGAGGGGCGCGAGCGCGGCTCAGCGCCATCCGCGAGCGCGCTGGGG 660				
QY 201 ProLeuVal 203				
Db 661 CCCCTGGTG 669				
RESULT 6				
AAS22437				
ID AAS22437 standard; cDNA: 1279 BP.				
AC AAS22437:				
XX 24-Oct-2001 (first entry)				
DE Human cDNA encoding a novel human protein #3.				
XX Human; novel protein; ss; Antinaemic; osteopathic; antiinflammatory;				
KW immunomodulatory; cytosolic; neuroprotective; vulnerability; neotopic;				
KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;				
KW antibacterial; antiallergic; dermatological; hemostatic; antisthmatic;				
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;				
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;				
XX tissue regeneration; immune disorder.				
XX Homo sapiens.				
XX				
PN W020015437-A2.				
XX				

PD 02-AUG-2001.
 XX 25-JAN-2001; 2001WO-US02623.
 XX 25-JAN-2000; 2000US-0491404.
 PR (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT;
 PI
 XX
 XX
 DR MPI: 2001-451939/48.
 DR P-PSDB; AA014132.
 PT Isolated polypeptides useful for treating anti-inflammatory diseases,
 PS nervous system disorders, and for regenerating bone and cartilage -
 Claim 1: Page 167-169; 894pp: English.

CC The invention relates to polynucleotides encoding novel human
 CC proteins or their active domains. The polypeptides, polynucleotides and
 CC antibodies raised against the polypeptides are used in a method of
 CC treatment of a mammal and prevention of disorders caused by the aberrant
 CC protein expression or activity. The polypeptides can be used as
 CC molecular weight markers, food supplements, and in antibody production.
 CC The polypeptides are used to identify compounds which bind to the
 CC polypeptides. Polynucleotides of the invention are used as probes and
 CC primers, for sequencing, for chromosome or gene mapping, in the
 CC production of recombinant proteins, and in generating anti-sense DNA or
 CC RNA and in gene therapy. Polypeptides of the invention can be used to
 CC target drugs to a tumor. In assays to determine biological activity, to
 CC raise antibodies/ elicit an immune response, to determine quantitative
 CC protein levels, as tissue markers, and to isolate receptors or ligands.
 CC Polypeptides of the invention may also be useful in treating platelet
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting
 CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,
 CC anti-inflammatory diseases, nervous system disorders, and infection.
 CC The present sequence encodes a protein of the invention.

XX
 XX
 SQ Sequence 1279 BP; 238 A; 402 C; 476 G; 163 T; 0 other;

Alignment Scores:
 Pred. No.: 1,68e-87 Length: 1279
 Score: 1014.00 Matches: 203
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x AAS22437 (1-1279)

QY 1 MetIysValleuTrpAlaalaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
 DB 159 ATGAAGTTCTGTGGCTGCTGTCGTCATCTCTGCGAGAGGCCAGGCGT 218
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnThrGluThrGlnSer 40
 DB 219 GAGCAAGCGGTGAGACAGAGCCGAGCCGACCTCGCCAGACAGCCGAGTGGCAGAGC 278
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB 279 GGCACAGCGCTGGGAACTGGCACTGGGTGCTTTGGGATTACTGCGCTGGGTCAGACA 338
 QY 61 LeuSerGluGlnAlaGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB 339 CTGTCTGAGAGAGGTGACAGAGAGTGTCTACGCTCCAGAGTACCCAGAACGTAGAGGCG 398
 QY 81 LeuMetAspGluThrMetIysGluLeuLysAlaTyrIysSerGluLeuGluGlnLeu 100

DB 399 CTGATGAGACGAGACCATGAAAGGAGTTGCAAGGCCCTACAAATCGGAACGTGAGGAACTG 458
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerIysGluLeuGlnAlaValGlnAla 120
 DB 459 ACCCGGTGGCGAGGAGAGCGCGGACCGCTCTCCAGAGAGCTGAGGCGGCGAGGCC 518
 QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 DB 519 CGGCTGGCGCGGACATGAGAGAGCGTGTGGCGCGCTGTGTCAGTACCGCGCGAGGTC 578
 QY 141 GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 579 CAGCCATGCTCGCGCAGAGACCGAGAGCTGCGGGTGGCTCCCTCCACCTGCGC 638
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 DB 639 AACCTGCGTAAAGCGGCTCTCCCGATGCCGATGACCTGCAGAAAGCGCTGGCAGTGTAC 698
 QY 181 GlnAlaGlyAlaArgGluGluGluAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 699 CAGCGCGGGGCGCGCAGAGGCGCGGAGCGCGGCTCAGCGCCATCCGCGAGCGCTGGGG 758
 QY 201 ProLeuVal 203
 DB 759 CCCCTGTGTG 767

RESULT 7
 AAD22052
 ID AAD22052 standard; DNA; 1156 BP.
 XX AAD22052;
 AC
 XX
 XX
 DT 12-FEB-2002 (first entry)
 XX
 DE Human apolipoprotein E (apoE) allele, apoE2** DNA.
 XX
 XX Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;
 KW hypertriglyceridaemia; low density lipoprotein; LDL; ds.
 XX
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 61..1014
 FT /tag= a
 FT /product= "Human apoE allele, apoE2**"
 FT sig_peptide 61..114
 FT /tag= b
 FT mat_peptide 115..1011
 FT /tag= c
 FT /product= "Mature human apoE allele, apoE2**"
 PN WO200177136-A1.
 XX
 PD 18-OCT-2001.
 XX
 XX
 PF 06-APR-2001; 2001WO-US11358.
 XX
 PR 06-APR-2000; 2000US-0544386.
 PR 04-OCT-2000; 2000US-0679088.
 PR 05-APR-2001; 2001US-0827854.
 XX
 PA (KOSP-) KOS PHARM INC.
 PA (UYBO-) UNIV BOSTON.
 PI Zannis VI, Kypreos KE;
 XX
 XX
 DR MPI: 2002-010885/01.
 DR P-PSDB; AAE13298.
 PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
 PT cholesterol, delaying the onset of or treating atherosclerosis in
 PT mammal, without inducing hypertriglyceridaemia -

XX Claim 14; Page 83; 91pp; English.

PS The present sequence is a human apolipoprotein E (apoE)

XX allele, apoE2** DNA. The apoE lipoproteins are useful for

CC lowering cholesterol, delaying the onset of atherosclerosis,

CC treating or regressing atherosclerosis without inducing

CC hypertriglyceridaemia, in a mammal lacking an endogenous,

CC normally functioning apoE gene or low density lipoprotein (LDL)

CC receptor or is at risk for developing atherosclerosis due to

CC accumulation of lipoprotein remnants in the bloodstream or having

CC a defect in remnant removal.

XX

SQ Sequence 1156 BP; 207 A; 369 C; 432 G; 148 T; 0 other;

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
3	58e-87	1010.00	100.00%	99.51%	1156	202	1	0	0	0

DB: 24

US-09-827-854-15_COPY_1_203 (1-203) x AAD22052 (1-1156)

OY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20

DB 61 ATGAAGGTTCTGTGGGCTCGTTGGTGCATTCCTGGCAGAGTCCAGGCGCAAGTG 120

OY 21 GlnGlnAlaValAlaGlnThrGluProGluProGluLeuArgGlnGlnThrGlnSer 40

DB 121 GAGCAAGCGGTGGAGACAGAGCGCGGAGCTGCGCAGACAGCCAGTGGCAGAGC 180

OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60

DB 181 GGCCAGCGTGGGAATGACACTGGGTGCTTTGGATTACCTGGCGGGGAGACA 240

OY 61 LeuSerGlnGlnValGlnGlnLeuLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80

DB 241 CTGTCTGACAGAGTGACAGAGAGAGCTGCTCAGCTCCAGTACCCAGAGAACTGAGGGCG 300

OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100

DB 301 CTGATGAGCAGAGACCATGAAAGAGTTGAAGGCTTACAAATCGAAGCTGAGGAACTG 360

OY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnGlnAlaGlnAla 120

DB 361 ACCCCGCTGGCGAGAGAGCGGGGCGGCTGTCMAAGAGCTGACAGCGGGCGAGGGCC 420

OY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140

DB 421 CGGCTGGGCGCGGACATGAGACGCTGCGGCGGCTGCTGATGACGAGCGGAGGTG 480

OY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160

DB 481 CAGGCGCATGCTGGCGCAGACCGGAGAGCTGGGGTGGCGCTGCCACCTCGCGCC 540

OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180

DB 541 AAGCTCGTACAGCGGCTCTCGCGGATGCGATGACCTCAAGAGGCGCTGGCAGAGTAC 600

OY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200

DB 601 CAGCGCGGGGCGCGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660

OY 201 ProLeuVal 203

DB 661 CCCCTGTG 669

RESULT 8

AAF84314

ID AAF84314 standard; cDNA; 1156 BP.

XX

AC AAF84314;

XX 21-JUN-2001 (first entry)

DT XX

XX Human ApoE2 coding sequence.

DE XX

XX Human; ApoE2; Alzheimer's disease; arteriosclerosis; ss.

KW XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 61..1014

FT /*tag= a

FT /product= "Human ApoE2"

XX

PN JP2001017028-A.

XX

PD 23-JAN-2001.

XX

XX 28-APR-2000; 2000JP-0128919.

PF

XX 06-MAY-1999; 99JP-0125647.

PR

XX (MITU) MITSUBISHI CHEM CORP.

PA

XX WPI; 2001-285406/30.

DR

XX P-PSDB; AAB80996.

DR

XX

PT New apoE humanized mammalian cell useful for screening for agents

PT useful for treating or preventing Alzheimer's disease and

PT arteriosclerosis -

PR

XX

PS Disclosure; Page 11-12; 22pp; Japanese.

XX

CC The present invention relates to an ApoE humanised mammalian cell. The

CC present sequence is the coding sequence for human ApoE2, which was used

CC in the method of the present invention. The ApoE humanised mammalian cell

CC can be used for screening for agents useful for treating or preventing

CC Alzheimer's disease and arteriosclerosis.

XX

SQ Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
8	62e-87	1006.00	99.51%	99.21%	1156	202	1	0	0	0

DB: 22

US-09-827-854-15_COPY_1_203 (1-203) x AAF84314 (1-1156)

OY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20

DB 61 ATGAAGGTTCTGTGGGCTCGTTGGTGCATTCCTGGCAGAGTCCAGGCGCAAGTG 120

OY 21 GlnGlnAlaValAlaGlnThrGluProGluProGluLeuArgGlnGlnThrGlnSer 40

DB 121 GAGCAAGCGGTGGAGACAGAGCGGAGCGGAGCTGCGCAGACAGCCAGTGGCAGAGC 180

OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60

DB 181 GGCCAGCGTGGGAATGACACTGGGTGCTTTGGATTACCTGGCGGGGAGACA 240

OY 61 LeuSerGlnGlnValGlnGlnLeuLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80

DB 241 CTGTCTGACAGAGTGACAGAGAGAGCTGCTCAGCTCCAGTACCCAGAGAACTGAGGGCG 300

OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnGlnGln 100

DB 301 CTGATGAGCAGAGACCATGAAAGAGTTGAAGGCTTACAAATCGAAGCTGAGGAACTG 360

OY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnGlnAlaGlnAla 120

|||||
Db ACCCGGTGGCGGAGAGACCGCGGCGACGGCTGTCCAAAGAGCTGCAGCGCGCGCAGGCC 420
Qy ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
Db 421 CGGCTGGCGCGGACATGGAGAGACGTGTGGCGCGCTGGTGGCATCCGGCGGAGGTG 480
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTCGGCGACAGACCGCAGAGACGTGGGGTGGCGCTCGCCCTCCACCTGCGC 540
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTGGGTAGCGGCTCTCCGCGATGCCGATGACCTGAGAGATGCTGGCGAGTAC 600
Qy 181 GlnAlaGlyAlaArgGlyGluGlyAlaGluArgGlyLeuSerAlaIleArgGlyArgLeuGly 200
Db 601 CAGGCCGGGCGCGCGAGGGCGCGGACCGCGCTCAAGCGCATCCGCGAGCGCTGGGG 660
Qy 201 ProLeuVal 203
Db 661 CCCCTGGTG 669
RESULT 9
AAD22049
ID AAD22049 standard; DNA: 1156 BP.
AC AAD22049;
XX
DT 12-FEB-2002 (first entry)
XX
DE Human apolipoprotein E (apoE) isoprotein, apoE2 DNA.
XX
KM Human: apolipoprotein E; apoE: cholesterol; atherosclerosis;
KM hypertriglyceridaemia; low density lipoprotein; LDL; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 61..1014
FT /*tag= a
FT /product= "Human apoE isoprotein, apoE2"
FT sig_peptide 61..114
FT /*tag= b
FT mat_peptide 115..1011
FT /*tag= c
FT /product= "Mature human apoE isoprotein, apoE2"
XX
PN MO200177136-A1.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001MO-US11358.
XX
PR 06-APR-2000; 2000US-0544386.
PR 04-OCT-2000; 2000US-0679088.
PR 05-APR-2001; 2001US-0827854.
XX
PA (KOSP-) KOS PHARM INC.
PA (UYBO-) UNIV BOSTON.
XX
PI Zannits VI, Kypros KE;
XX
DR MPI: 2002-010885/01.
DR P-PSDB: AAEL3295.
XX
PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT cholesterol, delaying the onset of or treating atherosclerosis in
PT mammal, without inducing hypertriglyceridaemia. -
XX
PS Claim 14; Page 81-82; 91pp; English.
XX
CC The present sequence is a human apolipoprotein E (apoE)

CC isoprotein, apoE2 DNA. The apoE lipoproteins are useful for
CC lowering cholesterol, delaying the onset of atherosclerosis,
CC treating or regressing atherosclerosis without inducing
CC hypertriglyceridaemia, in a mammal lacking an endogenous,
CC normally functioning apoE gene or low density lipoprotein (LDL)
CC receptor or is at risk for developing atherosclerosis due to
CC accumulation of lipoprotein remnants in the bloodstream or having
CC a defect in remnant removal.
XX
SQ Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;
Alignment Scores:
Pred. No.: 8,62e-87 Length: 1156
Score: 1006.00 Matches: 202
Percent Similarity: 99.51% Conservative: 0
Best Local Similarity: 99.51% Mismatches: 1
Query Match: 99.21% Indels: 0
DB: 24 Gaps: 0
US-09-827-854-15_COPY_1_203 (1-203) x AAD22049 (1-1156)
Qy 1 MetLysValLeuThrPalaIalaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGTCTCTGTGGCTGCTGCTGTGTCACATTCCTGCAGAGATGCCAGCCAGGTG 120
Qy 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAAGCGGTGGAGAGACAGACCGGAGCCGAGCTGGCGCAGAGACCGATGGCAGAGC 180
Qy 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 181 GGCCAGCGCTGGGAAACGCGACCTGGCTCTTGGATTACCTGGCTGGGTCAGACACA 240
Qy 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGTCTAGCAGAGTGGCAGAGAGCTGTCACTCCAGGTCAACCGAAGTGAAGGCGG 300
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
Db 301 CTGATGAGCAGACCATGAGAGAGATTGAAAGCTTACAAATCGGAACCTGAGAAACTG 360
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 361 ACCCGGTGGCGGAGAGACGCGGCGCGCTGTCCAAAGAGCTGACAGCGCGCAGGCC 420
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
Db 421 CGGCTGGCGCGGACATGGAGAGACGTGTGGCGCGCTGGTGGCATCCGGCGGAGGTG 480
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTCGGCGACAGACCGCAGAGACGTGGGGTGGCGCTCGCCCTCCACCTGCGC 540
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTGGGTAGCGGCTCTCCGCGATGCCGATGACCTGAGAGATGCTGGCGAGTAC 600
Qy 181 GlnAlaGlyAlaArgGlyGluGlyAlaGluArgGlyLeuSerAlaIleArgGlyArgLeuGly 200
Db 601 CAGGCCGGGCGCGCGAGGGCGCGGACCGCGCTCAAGCGCATCCGCGAGCGCTGGGG 660
Qy 201 ProLeuVal 203
Db 661 CCCCTGGTG 669
RESULT 10
AAD22051
ID AAD22051 standard; DNA: 1156 BP.
XX
AC AAD22051;
XX
DT 12-FEB-2002 (first entry)
XX

DE Human apolipoprotein E (apoE) allele, apoE2* DNA.
 XX
 KW Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;
 KM hypertriglyceridaemia; low density lipoprotein; LDL; ds.
 XX
 OS Homo sapiens.
 XX
 XX
 XX Key Location/Qualifiers
 FT CDS 61..1014
 FT /tag= a
 FT /product= "Human apoE allele, apoE2*" 61..114
 FT /tag= b
 FT mat_peptide 115..1011
 FT /tag= c
 FT /product= "Mature human apoE allele, apoE2*" 115..1011
 XX
 XX W0200177136-A1.
 XX
 XX 18-OCT-2001.
 XX
 XX 06-APR-2001; 2001WO-US11358.
 XX
 XX 06-APR-2000; 2000US-0544386.
 PR 04-OCT-2000; 2000US-0679088.
 PR 05-APR-2001; 2001US-0827854.
 XX
 PA (KOSP-) KOS PHARM INC.
 PA (UYBO-) UNITIV BOSTON.
 XX
 PI Zannis VI, Kypreos KE;
 XX
 DR WPI: 2002-010885/01.
 DR P-PSDB; AAE13297.
 XX
 PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
 PT cholesterol, delaying the onset of or treating atherosclerosis in
 PT mammal, without inducing hypertriglyceridaemia -
 XX
 PS Claim 14; Page 82; 91pp: English.
 XX
 XX The present sequence is a human apolipoprotein E (apoE)
 CC allele, apoE2* DNA. The apoE lipoproteins are useful for
 CC lowering cholesterol, delaying the onset of atherosclerosis,
 CC treating or regressing atherosclerosis without inducing
 CC hypertriglyceridaemia, in a mammal lacking an endogenous,
 CC normally functioning apoE gene or low density lipoprotein (LDL)
 CC receptor or is at risk for developing atherosclerosis due to
 CC accumulation of lipoprotein remnants in the bloodstream or having
 CC a defect in remnant removal.
 XX
 XX Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other:
 SQ

Alignment Scores:
 Pred. No.: 8.62e-87 Length: 1156
 Score: 1006.00 Matches: 202
 Percent Similarity: 99.51% Conservative: 0
 Best Local Similarity: 99.51% Mismatches: 1
 Query Match: 99.21% Indels: 0
 DB: 24 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x AAD22051 (1-1156)

OY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 61 ATGAAGGTTCTGTGGCTGCTGCTGTCACATTCTGGCAGATGCCAGCCAAAGGTG 120
 OY 21 GlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 121 GAGCAAGCGGTGAGACACAGCCGAGCCGAGCTGGCCACACAGCAGTGGCAGAGC 180
 OY 41 GlnGlnArgTrpGluLeuAlaLeuGlnLysArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 181 GGCCAGCGCTGGGAACCTGGACACTGGGTGCTTTTGGAGTTACTGCGCTGGGTGACAGACA 240
 OY 61 LeuSerGluGlnValGluGluGluLeuLeuSerSerGlnValThrGluGluLeuArgAla 80
 DB 241 CTGTCTGAGCAGGTGCTGAGAGAGAGCTGCTCAGCTCCAGGTCACCCAGAACTGAGGCG 300
 OY 81 LeuMetAspGluTrpMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu 100
 DB 301 CTGATGAGACGAGACCATGATGAGAGCTTGAAGGCTTCAATCGAATCGAGACTGGAGAACACTG 360
 OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
 DB 361 ACCCGGTGGCGGAGAGACGCGGCGACGCTGTCCAAGAGCTGCAGGCGCGCAGGCGC 420
 OY 121 ArgLeuGlyAlaAspMetGlnAspValLysGlyArgLeuValGlnTyrArgGlyGlyVal 140
 DB 421 CGGCTGGCGCGGACATGAGAGAGCTGTGGCGCGCTGTGTCACTATCCGCGGCAAGGTG 480
 OY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 481 CAGGCCATCTCTGGCGCAGACGACCGAGAGCTGGCGGTGCGCTGCTCCACCTGGCGC 540
 OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 DB 541 AAGCTGTGAAGCGGCTCTCTCGCGATGCGCTGACCTGACAGAAAGCGCTGGCAGTAC 600
 OY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 601 CAGCGCGGGCGCGCGAGGGCGCGAGCGCGCGCTCAGCGCATCCGAGGAGCGCTGGGG 660
 OY 201 ProLeuVal 203
 DB 661 CCCCTGGTG 669

RESULT 11
 AAT06957
 ID AAT06957 standard; cDNA to mRNA; 1157 BP.
 XX
 AC AAT06957;
 XX
 DT 19-JUN-1996 (first entry)
 XX
 DE Human apolipoprotein-E (ApoE) cDNA.
 XX
 KW Recombinant; human; apolipoprotein-E; ApoE; insect cells; larva;
 KW Manduca sexta; Autographa californica nuclear polydrosis virus;
 KW haemolymph; lipid complex; biologically active; ds.
 XX
 OS Homo sapiens.
 XX
 XX
 XX Key Location/Qualifiers
 FT MISC-feature 62..119
 FT /tag= a
 FT note= "misc_signal" 62..1015
 FT CDS /tag= b
 PN US5472858-A.
 XX
 XX 05-DEC-1995.
 XX
 XX 04-JUN-1991; 91US-0709949.
 XX
 XX 04-JUN-1991; 91US-0709949.
 PR 04-JUN-1991; 91US-0709949.
 XX
 XX (MISC) WISCONSIN ALUMNI RES FOUND.
 PI Attie AD, Beckage NE, Gretch DG, Sturley SL.
 XX
 DR WPI: 1996-029812/03.
 DR P-PSDB; AAR86791.
 XX
 PT Prodn. of recombinant apo:lipoprotein E in insects - by infecting

QY 201 ProLeuVal 203
| | | | |
XX 662 CCCCTGGTG 670
Db
RESULT 14
ABL5450
ID ABL65450 standard; DNA: 1157 BP.
XX
AC ABL65450;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:3787.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KM stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KM cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN W0200194629-A2.
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 18-SEP-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 20-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 22-SEP-2000; 2000US-234053P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237279P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX

PI Young PE, Augustus M, Carter KC, Edner R, Endress G, Horrigan S;
PI Sopfel DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1; SEQ ID 3787; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;
XX
Alignment Scores:
Pred. No.: 1,07e-86 Length: 1157
Score: 1005.00 Matches: 201
Percent Similarity: 99.01% Conservative: 0
Best Local Similarity: 99.01% Mismatches: 2
Query Match: 99.11% Indels: 0
DB: Gaps: 24
US-09-827-854-15_COPY_1_203 (1-203) x ABL65450 (1-1157)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db ATGAAAGGCTTCGTGGCGGTGGCTGGCTGCACATTCCTGCAGAGATCCAGGCCAAGGTG 121
QY 21 GluGlnAlaValAlaGlnTrpGlnProGlnProGlnLeuArgGlnGlnTrpGlnSer 40
Db GAGCAAGCGGTGCAGACAGAGCCGAGCCGAGCTGCGCCAGACAGCCAGTGGCAGAGC 181
QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db GCCCAGCGGTGGAACTGGCACTGGTCGCTTTGGATTACTGCGCTGGTGGAGACA 241
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerGlnValThrGlnGlnLeuArgAla 80
Db CTGTCTGAGCAGAGGTGCAGAGAGCTGCTCAGCTCCCAATGCCAACAAATGAGGGCG 301
QY 242 CTGTCTGAGCAGAGGTGCAGAGAGCTGCTCAGCTCCCAATGCCAACAAATGAGGGCG 301
QY 81 LeuMetAspGlnTrpMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnGln 100
Db CTGATGACAGACCATGAAGAGATTGAAGCCTTACAAATCGAATCGAGAGAACTG 361
QY 101 ThrProValAlaGlnGlnTrpArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
Db ACCCGGTAGCGAGAGAGAGCGGGGCGCGCTGTCCAAAGAGCTGCAGAGCGCGAGGCC 421
QY 362 ACCCGGTAGCGAGAGAGAGCGGGGCGCGCTGTCCAAAGAGCTGCAGAGCGCGAGGCC 421
QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnTrpArgGlnGlnVal 140
Db CGGCTGGCGCGGACATGAGAGACGTGCGCGCGCTGTGCAGTACCCCGCGAGGTG 481
QY 141 GlnAlaMetLeuGlnGlnSerTrpGlnGlnGlnArgValArgLeuAlaSerHisLeuArg 160
Db CAGGCCATGCTCGCCAGACAGACCGAGAGAGCTCGGGGTGGCTCGCTCCACCTGCGC 541

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaaspAspLeuGlnLysArgLeuAlaValTyr 180
 DB 542 AAGCTGCGGTAAAGCGGCTCTCCGCGCATGCCATGACAGCAAGCGCCTGGCAGTGTAC 601
 QY 181 GlnAlaGlyAlaArgGlnGlyAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
 DB 602 CAGGCGGGGGCCCGGAGGCGCGCGAGCGCGCTCAGCGCCATCGGAGCGCCTGGGG 661
 QY 201 ProLeuVal 203
 DB 662 CCCCTGGTG 670
 RESULT 15
 AAN50450
 ID AAN50450 standard; DNA; 1110 BP.
 XX
 AC AAN50450;
 XX
 DT 09-JAN-1992 (first entry)
 XX
 DE Sequence encoding human apolipoprotein E.
 XX
 KM Hyperlipaemia; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 15..968
 FT /tag= a
 FT mat_peptide 69..965
 FT /*tag= b
 XX
 PN JP60118189-A.
 XX
 PD 25-JUN-1985.
 XX
 PF 29-NOV-1983; 83JP-0224980.
 XX
 PR 29-NOV-1983; 83JP-0224980.
 XX
 PA (MITU) MITSUBISHI CHEM IND KK.
 XX
 DR WPI; 1985-188003/31.
 DR P-PSDB; AAP51204.
 XX
 PT DNA fragment - contg. DNA which codes human apolipoprotein E for
 PT treatment of hyperlipaemia.
 XX
 PS Claim 3; Page 484; 8pp; Japanese.
 CC The sequence may be used to produce the apolipoprotein E, useful in the
 CC treatment of hyperlipaemia.
 XX
 SQ Sequence 1110 BP; 198 A; 355 C; 414 G; 143 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,27e-86 Length: 1110
 Score: 1004.00 Matches: 201
 Percent Similarity: 99.01% Conservative: 0
 Best Local Similarity: 99.01% Mismatches: 2
 Query Match: 99.01% Indels: 0
 DB: 6 Gaps: 0
 US-09-827-854-15_COPY_1_203.(1-203) x AAN50450 (1-1110)
 QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 15 ATGAAGGTTCTGTGGCTGCTGCTGGTCAATTCCTGGCAGATGCCAGGCCAAAGTG 74
 QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnTrpGlnSer 40
 DB 75 GAGCAAGCGGTGGAGACAGAGCCGAGCGCGGACAGACAGACCGAGTGGCAGAGAC 134

QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlnLysArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB 135 GGGCAGCGGTGGGAACCTGGACATGGGGTGGCTTTTGGATTACTGCGGTGGGTGAGACA 194
 QY 61 LeuSerGlnGlnValGlnGlnLeuLeuLeuSerSerGlnValThrGlnLeuArgAla 80
 DB 195 CTGTCTGAGCACGTGCAGAGAGAGCTGTCACTCCACAGTCCAGAGAACCTGAGGGCG 254
 QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
 DB 255 CTGATGAGACGAGACCATGAGAGGTTGAGAGGCTTCAAAATCGGAACCTGAGGAGCAACTG 314
 QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
 DB 315 ACCCGGTGGCGGAGAGAGAGCGGGGACCGCTGTCCAAAGACCTGCAGCGCGCCAGGCC 374
 QY 121 ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
 DB 375 CGGCTGGGGCGGACATGAGAGACGTGTGCGCGCCCTGTGTCAGTACCGCGGAGGTG 434
 QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 435 CAGGCCATGCTCGGCCAGAGACCGAGAGAGCTGGGGGTGGCTGCCTGCCACCTGGCC 494
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaaspAspLeuGlnLysArgLeuAlaValTyr 180
 DB 495 AAGCTGCGGTAAAGCGGCTCTCCGCGCATGCCATGACCTGCAGCAAGCGCCTGGCAGTGTAC 554
 QY 181 GlnAlaGlyAlaArgGlnGlyAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
 DB 555 CAGGCGGGGGCCCGGAGGCGCGCGAGCGCGCTCAGCGCCATCCGCGAGCGCCTGGGG 614
 QY 201 ProLeuVal 203
 DB 615 CCCCTGGTG 623

Search completed: March 14, 2003, 12:43:05
 Job time : 117.454 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:17:52 ; Search time 22.2431 Seconds
(without alignments)
2798.866 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_203

Perfect score: 1014

Sequence: 1 MKVLMALVTLFGLGCAKV.....AREGAEGLSIRRLGPLY 203

Scoring table:

BLASTSUM62	
Xgapop 10.0 , Xgapext 0.5	
Xgapop 10.0 , Ygapext 0.5	
Xgapop 6.0 , Ygapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus.p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09827854/runat_11032003_101611_27510/app_query.fasta_1.3576
-DB=Issued_Patents_NA -QPMT=fastapi -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bloms62 -TRANS=human40.cdl
-LIST=45 -DOCCALIGN=200 -THR SCORE=100 -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORR=ext -HEAPSIZE=500 -MITLEN=0 -MAXLEN=200000000
-USER=US09827854.ecgn_1.1-224-etunat_11032003_101611_27510 -NCPV=6 -ICPV=3
-NO_XLPHY -NO_MMAP -LARGESUBSTR -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA:
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCtus.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1005	99.1	1157	1	US-07-709-949-1
2	984	97.0	660	2	US-08-726-306A-28
3	724.5	71.4	1126	4	US-08-949-155-5
4	724.5	71.4	1126	4	US-09-819-964-5
5	588.5	58.0	4267	4	US-08-949-155-51
6	588.5	58.0	4267	4	US-09-819-964-51
7	511	50.4	330	1	US-07-849-389-6
8	374	36.9	252	3	US-08-617-256-24
9	374	36.9	252	4	US-09-287-141-24
10	374	36.9	252	4	US-09-431-613-24
11	374	36.9	252	4	US-09-504-245-24
12	374	36.9	252	4	US-09-287-682-24

13	374	36.9	252	4	US-09-287-679-24
14	374	36.9	252	4	US-09-397-766-24
15	374	36.9	252	4	US-09-287-681-24
16	374	36.9	252	4	US-09-495-444-24
17	132	13.0	801	1	US-07-959-946-4
18	132	13.0	801	1	US-08-333-577-4
19	132	13.0	801	5	PCT-US92-08634-4
20	132	13.0	842	4	US-08-952-796-1
21	123	12.1	516	4	US-09-183-861-34
22	123	12.1	516	4	US-09-022-765-34
23	123	12.1	11236	4	US-07-853-913-1
24	119	11.7	5661	4	US-08-938-105-2
25	118	11.6	592	4	US-09-276-531-115
26	118	11.6	603	4	US-08-952-796-14
27	118	11.6	1879	4	US-09-750-580-12
28	116	11.4	863	1	US-08-448-606-7
29	116	11.4	964	1	US-08-448-606-5
30	116	11.4	4296	4	US-09-060-410-3
31	112.5	11.1	3256	2	US-08-968-751-3
32	112	11.0	9551	1	US-08-056-200-93
33	112	11.0	9551	2	US-08-800-644-93
34	106	10.5	8789	1	US-08-328-254-5
35	106	10.5	10136	1	US-08-353-700-2
36	106	10.5	10136	5	PCT-US95-16216-2
37	105.5	10.4	5381	4	US-09-750-580-4
38	105.5	10.4	81001	4	US-09-750-580-1
39	105	10.4	4868	1	US-08-139-937-12
40	105	10.4	4868	5	PCT-US93-11310-12
41	103.5	10.2	6306	1	US-08-195-487-3
42	103.5	10.2	6306	5	PCT-US93-06160-3
43	102.5	10.1	3033	1	US-08-095-737-3
44	102.5	10.1	3033	1	US-08-480-145-3
45	102.5	10.1	3033	2	US-08-477-389-3

ALIGNMENTS

RESULT 1
US-07-709-949-1
Sequence 1, Application US/07709949
Patent No. 5472858
GENERAL INFORMATION:
APPLICANT: Attie, Alan D
APPLICANT: Gretch, Daniel G
APPLICANT: Sturley, Stephen L
APPLICANT: Beckage, Nancy E
TITLE OF INVENTION: Production of Recombinant Proteins in
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Quarles & Brady
STREET: P.O. Box 2113
CITY: Madison
STATE: WI
COUNTRY: USA
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/709, 949
FILING DATE: 19910604
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 9629691801
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 1:

Sequence 24, Appl
Sequence 24, Appl
Sequence 24, Appl
Sequence 24, Appl
Sequence 4, Appl1
Sequence 4, Appl1
Sequence 4, Appl1
Sequence 34, Appl
Sequence 34, Appl
Sequence 34, Appl
Sequence 115, App
Sequence 14, Appl
Sequence 7, Appl1
Sequence 5, Appl1
Sequence 3, Appl1
Sequence 3, Appl1
Sequence 93, Appl
Sequence 5, Appl1
Sequence 2, Appl1
Sequence 12, Appl
Sequence 3, Appl1
Sequence 3, Appl1
Sequence 3, Appl1
Sequence 3, Appl1

SEQUENCE CHARACTERISTICS:
LENGTH: 1157 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 62..1015
FEATURE:
NAME/KEY: misc.signal
LOCATION: 62..119
PUBLICATION INFORMATION:
AUTHORS: McLean, J W
JOURNAL: Journal of Biological Chemistry
VOLUME: 259
PAGES: 6498-6504
DATE: 1984
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 1157
US-07-709-949-1
Alignment Scores:
Pred. No.: 2.26e-105 Length: 1157
Score: 1005.00 Matches: 201
Percent Similarity: 99.01% Conservative: 0
Best Local Similarity: 99.01% Mismatches: 2
Query Match: 99.11% Indels: 0
Gaps: 0
US-09-827-854-15_COPY_1_203 (1-203) x US-07-709-949-1 (1-1157)
QY 1 MetlyValleuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 62 ATGAAGTTCTGTGGCTGTGCTGTCACATTCTTGGCAGAGATCCAGGCCAAGGTG 121
QY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluThr 40
DB 122 GAGCAAGCGGTGAGAGACAGCCGAGCCGAGCTGGCCAGCCAGACGATGGACAGC 181
QY 41 GlnGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 182 GGCAGGCGTGGAGACGGCAGCTGGCTTTGGGATTACCTGGCTGGTGCACACA 241
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerGlnValThrGlnGluLeuArgAla 80
DB 242 CTGCTGTGAGCAGGTGCGAGGAGCTGCTCAGCTCCCAAGTCAACCAAGTGAAGGCG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
DB 302 CTGATGACAGACACCATGAAGAGCTTAAGGCTTACAAATCGGAGACGAGAACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 362 ACCCGGTACGGAGAGAGCGGGGACGCTGCCAAGAGCGTGCAGAGCGGCGCAAGCC 421
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
DB 422 CGGCTGGCGCCGACATGAGGAGCTGTGCGCGCTGCTGACGATCCCGCGGAGAGTG 481
QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 482 CAGGCCCTGTGCGCCAGACACCGAGAGCTGCGGTGGCGCTCCGCTCCCACTGCGC 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 542 AAGCTGGGTAAGCGGCTCTCCCGGATCCCGATGACCTGAGAAAGCCCTGGCAGTAC 601
QY 181 GlnAlaGlyAlaArgGluGlnAlaGluArgGlyLeuSerAlaLeuArgGluArgLeuGly 200

DB 602 CAGCCCGGGCGCCGAGGCGCGGAGCGGCGCTCAGCCCATCCGAGCGCTGCGG 661
QY 201 ProLeuVal 203
DB 662 CCCCTGGTG 670
RESULT 2
US-08-726-306A-28
Sequence 28, Application US/08726306A
Patent No. 5958684
GENERAL INFORMATION:
APPLICANT: van Leeuwen, Frederik Willem
APPLICANT: Burbach, Johannes Peter Henri
APPLICANT: Grosveld, Franklin G.
TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
NUMBER OF SEQUENCES: 189
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Wilcoff, Ltd.
STREET: 1 Financial Center
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,306A
FILING DATE: 02-Oct-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 95/20080.4
FILING DATE: 02-Oct-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/009,832
FILING DATE: 01-Jan-1996
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Ph.D., Kathleen M.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 345-9100
TELEFAX: (617) 345-9111
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-726-306A-28
Alignment Scores:
Pred. No.: 2.51e-103 Length: 660
Score: 984.00 Matches: 197
Percent Similarity: 98.99% Conservative: 0
Best Local Similarity: 98.99% Mismatches: 2
Query Match: 97.04% Indels: 0
Gaps: 0
US-09-827-854-15_COPY_1_203 (1-203) x US-08-726-306A-28 (1-660)
QY 1 MetlyValleuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 62 ATGAAGTTCTGTGGCTGTGCTGTCACATTCTTGGCAGAGATCCAGGCCAAGGTG 121
QY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluThr 40
DB 122 GAGCAAGCGGTGAGAGACAGCCGAGCCGAGCTGGCCAGCCAGACGATGGACAGC 181
QY 41 GlnGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60

|||||
Db 182 GCCAGCGCTGGAACTGGACACTGGGCTTTGGATTACCTGGCGCTGGTGCAGACA 241
Oy 61 LeuSerGluGlnValGlnGluGlnLeuSerSerGlnValThrGlnGlnLeuAlaGla 80
Db 242 CTGTCTAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 301
Oy 81 LeuMetAspGluThrMetGluLeuValGluValGluValGluValGluValGluVal 100
Db 302 CTGATGACACAGACATGAAGAGGAGTGAAGGCTTACAAATCGGAGACTGGAGAACAACTG 361
Oy 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 362 ACCCGGTACCGGAGAGACGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 421
Oy 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnThrArgGlyGluVal 140
Db 422 CGGCTGGGCGGACATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 481
Oy 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
Db 482 CAGGCGATGCTGGCGGAGACACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 541
Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AAGCTGGTAAAGGCGCTCTCCGCGATCCCGATCCGATCCGATCCGATCCGATCCGAT 601
Oy 181 GlnAlaGlnAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 199
Db 602 CAGGCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 658

RESULT 3

US-08-949-155-5
Sequence 5, Application US/08949155

Patent No. 6271436

GENERAL INFORMATION:

APPLICANT: Piedrahita, Jorge A

APPLICANT: Bazer, Fuller W

TITLE OF INVENTION: Compositions and Methods for the

TITLE OF INVENTION: Generation of Transgenic Animal Species

NUMBER OF SEQUENCES: 51

CORRESPONDENCE ADDRESS:

ADDRESSEE: ARNOLD, WHITE AND DURKEE

STREET: P.O. Box 4433

CITY: Houston

STATE: TX

COUNTRY: US

ZIP: 77210-4433

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/949,155

FILING DATE: Concurrently Herewith

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/027,338

FILING DATE: 11-OCT-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/046,094

FILING DATE: 09-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Hibler, David W.

REGISTRATION NUMBER: 41,071

REFERENCE/DOCKET NUMBER: TANK:177

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (713) 789-2679

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1126 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE: CDS
NAME/KEY: CDS
LOCATION: 51..1001
US-08-949-155-5

Alignment Scores:
Pred. No.: 1,9e-73 Length: 1126
Score: 724.50 Matches: 146
Percent Similarity: 82.52% Conservative: 24
Best Local Similarity: 70.87% Mismatches: 29
Query Match: 71.45% Indels: 7
DB: 4 Gaps: 2

US-09-827-854-15_COPY_1_203 (1-203) x US-08-949-155-5 (1-1126)

Oy 1 MetLysValLeuThrPalaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 51 ATGAGGCTTGTGGGTGCTTTGGTGTACCTCTCCAGAGATCCGACAGAGAGAC 110
Oy 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArg-----GlnGlnThrGln 37
Db 111 GAGCGGGG-----CCGCGCGGAGGAGTGCACGTGTGTGGAGAGGCCAAG 158
Oy 38 TrpGlnSerGlyGlnArgTrpGlnLeuAlaGlnGlyArgPheTrpAspTyrLeuArgTrp 57
Db 159 TGGCAGGGCAGCCAGCCCTGGGAGCAGGCGCTGGCGCTTGGGATTAACCTGGCTGG 218
Oy 58 ValGlnThrLeuSerGlnGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 77
Db 219 GTTCAGTCCCTGTCTGTACCAAGTGCAGAGAGAGCTGTCAAGCAACAGTCAACCAAGAA 278
Oy 78 LeuArgAlaLeuMetAspGlnThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlu 97
Db 279 CTGACGAGCTGATAGAGAGACATGAAGAGGTGAAGGCTTAACCGCAGAGAGCTGGAG 338
Oy 98 GlnGlnLeuThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGluLeuGlnAla 117
Db 339 GCGCAGCTGGGCGCCGACCCAGCAGAGAGCAGGCGGCGCTGTCCAAAGAGCTGCAGGCG 398
Oy 118 AlaGlnAlaArgLeuGlnGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArg 137
Db 399 GCGCAGGCGCGCGTGGCGCGCGCAGATGAGAGAGCTGCGCAACCGCTTGGTGTCTACCC 458
Oy 138 GlyGluValGlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSer 157
Db 459 AGCGAGGTGCACAACATGTTGGGCGCAGACACCGAGAGGCTGCGGAGCCGCTGGCTTC 518
Oy 158 HisLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAspLeuGlnLysArgLeu 177
Db 519 CACCTCGCAGAGTGGCGCAGGCGGCTGCGGAGACCCAGGAGACCTGCGAAGAGCGCTG 578
Oy 178 AlaValTyrGlnAlaGlnAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 197
Db 579 GCGGTGTACAGGCGGCGGCGGCGGAGGCGCGGAGCGCGAGCGTGAAGCGCTTCCGCGAG 638
Oy 198 ArgLeuGlyProLeuVal 203
Db 639 CGCCTCGGCGCGCTGGTG 656

RESULT 4

US-09-819-964-5

Sequence 5, Application US/09819964

Patent No. 6369294

GENERAL INFORMATION:

APPLICANT: Piedrahita, Jorge A

APPLICANT: Bazer, Fuller W

TITLE OF INVENTION: Compositions and Methods for the

TITLE OF INVENTION: Generation of Transgenic Animal Species

NUMBER OF SEQUENCES: 51

CORRESPONDENCE ADDRESS:

ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/617,256
FILING DATE: March 18, 1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: March 17, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth A.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: SOL-013CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-617-256-24

Alignment Scores:
Pred. No.: 1,86e-34
Score: 374.00
Percent Similarity: 97.62%
Best Local Similarity: 97.62%
Query Match: 36.88%
DB: 3
Length: 252
Matches: 82
Conservative: 0
Mismatches: 2
Indels: 1
Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-08-617-256-24 (1-252)

QY 109 AlaArgLeuSerIysGluLeuGlnAlaIaGlnAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGTCGACGCGCGCGCGCTGGCGGCACATGGAGGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GGTGTGGC-CGCTGTGTGAGTACCGCGCGAGGTCACAGCCATGCTCGGCCAGAGCAC 120
QY 149 GluGlnLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTCGGGGTGGCTCCCTCCACCTGCGCAAGCTGCGTAAAGCGGCTCTCCGCC 180
QY 169 AspAlaAspLeuGlnIysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGTCCCTGCGAGTACCAAGCGCGGCGCGCGAGGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGGGCGCTC 252

RESULT 9
US-09-287-141-24
Sequence 24, Application US/09287141
Patent No. 6197498
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSER: Heller Ehrman White & McCaulliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA

ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,141
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-287-141-24

Alignment Scores:
Pred. No.: 1,86e-34
Score: 374.00
Percent Similarity: 97.62%
Best Local Similarity: 97.62%
Query Match: 36.88%
DB: 4
Length: 252
Matches: 82
Conservative: 0
Mismatches: 2
Indels: 1
Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-287-141-24 (1-252)

QY 109 AlaArgLeuSerIysGluLeuGlnAlaIaGlnAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGTCGACGCGCGCGCGCTGGCGGCACATGGAGGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GGTGTGGC-CGCTGTGTGAGTACCGCGCGAGGTCACAGCCATGCTCGGCCAGAGCAC 120
QY 149 GluGlnLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTCGGGGTGGCTCCCTCCACCTGCGCAAGCTGCGTAAAGCGGCTCTCCGCC 180
QY 169 AspAlaAspLeuGlnIysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGTCCCTGCGAGTACCAAGCGCGGCGCGCGAGGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGGGCGCTC 252

RESULT 10
US-09-431-613-24
Sequence 24, Application US/09431613
Patent No. 6221601
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33


```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/431,613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002G
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-587-5360
TELEFAX: 619-450-8400
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-431-613-24

Alignment Scores:
Pred. No.: 1,86e-34 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 36.88% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-431-613-24 (1-252)
QY 109 AlAArgLeuSerLySGluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
DB 2 GCACGGCTGTCCAAAGAGCTGCAGGGCGCCAGCCCGCTGGGGCGGAGCATGGAGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGGCC-GGCTGTGGCTGACCGCGGCGAGGTGCAGGCATGTCTCGCCGAGACACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGGGGGTGCCCTCGCTCCACCTCGCGCAAGCTGGCTAAGCGGCTCTCCGC 180
QY 169 AspAlaAspAspLeuGlnTyrArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGAGTCCCTGGCAGTGTACACAGCGCGGGCCCGCGAGGGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTCT 252
RESULT 11
US-09-504-245-24
```

```

Sequence 24, Application US/09504245
Patent No. 6221605
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe LLP
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/504,245
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002G
TELECOMMUNICATION INFORMATION:
TELEPHONE: 858-450-8400
TELEFAX: 858-587-5360
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-504-245-24

Alignment Scores:
Pred. No.: 1,86e-34 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 36.88% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-504-245-24 (1-252)
QY 109 AlAArgLeuSerLySGluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
DB 2 GCACGGCTGTCCAAAGAGCTGCAGGGCGCCAGCCCGCTGGGGCGGAGCATGGAGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGGCC-GGCTGTGGCTGACCGCGGCGAGGTGCAGGCATGTCTCGCCGAGACACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGGGGGTGCCCTCGCTCCACCTCGCGCAAGCTGGCTAAGCGGCTCTCCGC 180
QY 169 AspAlaAspAspLeuGlnTyrArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGAGTCCCTGGCAGTGTACACAGCGCGGGCCCGCGAGGGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTCT 252
```


RESULT 12
US-09-287-682-24
; Sequence 24, Application US/09287682
; Patent No. 6235478
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; APPLICANT: Little, Daniel P.
; APPLICANT: Braun, Andreas
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuiffie
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037-9103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/287,682
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 18-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/406,199
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-2002E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-587-5360
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-287-682-24

Alignment Scores:
Pred. No.: 1,86e-34 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 36.88% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-287-682-24 (1-252)
QY 109 AAlaArgLeuSerIysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGGGCGCGCCGCTGGCGCGGACATGCGAGC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGGCC-CGGCTGTGTGAGTACCGCGCGGCGAGGTGCAGGCCATGCTGCGCCAGAGCACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgGlyLeuArgGlySerArgLeuArg 168
DB 121 GAGAGACTCGGGGTGGCTCGCTCCACCTGCGCAGAGCTCGTAAAGCGGCTCTCCGCC 180

QY 169 AspAlaAspLeuGlnIlysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACTGCAGAAATCCCTGGCAGTGTACCAAGCGCGGCCCGCGAGGGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTCC 252

RESULT 13
US-09-287-679-24
; Sequence 24, Application US/09287679
; Patent No. 6258538
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; APPLICANT: Little, Daniel P.
; APPLICANT: Braun, Andreas
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuiffie
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037-9103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/287,679
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 18-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/406,199
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-2002C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-587-5360
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-287-679-24

Alignment Scores:
Pred. No.: 1,86e-34 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 36.88% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-287-679-24 (1-252)
QY 109 AAlaArgLeuSerIysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGGGCGCGCCGCTGGCGCGGACATGCGAGC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
| | | | |
Db 62 GTGTGGCC-CGCTGTGTGAGTACCGCGGAGGTGCAGGCATGTCTCGGCAGAGCACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
| | | | |
Db 121 GAGAGCTGGGGTGGCCCTCGCTCCACCTCGGCAAGCTGTAGCGGCTCTCCGC 180
QY 169 AspAlaAspAspLeuGlnTyrArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
| | | | |
Db 181 GATCCGATACCTGCAGAAAGTCCCTGTGACGTACAGGCCGGGGCCCGCAGAGGCGCC 240
QY 189 GluArgGlyLeu 192
| | | | |
Db 241 GAGCGCGGCTC 252
RESULT 14
US-09-397-766-24
; Sequence 24, Application US/09397766
; Patent No. 6268144
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037-9103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/397,766
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 18-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/406,199
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-20021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-587-5360
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-397-766-24
Alignment Scores:
Pred. No.: 1.86e-34 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 36.88% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-397-766-24 (1-252)
QY 109 AlaArgLeuSerLysGluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGluVal 128
| | | | |
Db 2 GCACGGCTGTCCAAAGAGCTGCAGCGCGCCGAGCCCGCTGGCGGACATGAGAGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
| | | | |
Db 62 GTGTGGCC-CGCTGTGTGAGTACCGCGGAGGTGCAGGCATGTCTCGGCAGAGCACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
| | | | |
Db 121 GAGAGCTGGGGTGGCCCTCGCTCCACCTCGGCAAGCTGTAGCGGCTCTCCGC 180
QY 169 AspAlaAspAspLeuGlnTyrArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
| | | | |
Db 181 GATCCGATACCTGCAGAAAGTCCCTGTGACGTACAGGCCGGGGCCCGCAGAGGCGCC 240
QY 189 GluArgGlyLeu 192
| | | | |
Db 241 GAGCGCGGCTC 252
RESULT 15
US-09-287-681-24
; Sequence 24, Application US/09287681
; Patent No. 6277573
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037-9103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/287,681
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 18-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/406,199
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-2002F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-587-5360
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-287-681-24
Alignment Scores:
Pred. No.: 1.86e-34 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 36.88% Indels: 1
DB: 4 Gaps: 0

Pred. No.: 1.86e-34
Score: 374.00
Percent Similarity: 97.62%
Best Local Similarity: 97.62%
Query Match: 36.88%
DB: 4

Length: 252
Matches: 82
Conservative: 0
Mismatches: 2
Indels: 1
Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-287-681-24 (1-252)

QY 109 AIAArgLeuSerLysGluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
Db 2 GCACGGCTGTCCAGAGAGCTGCAGGCGCGCCGCTGGCGCGGACATGGAGGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
Db 62 GTGTGCGC-CGCTGTGTGTCAGTACCGCGCGAGGTGCAGGCCATGCTCGGCCAGACACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
Db 121 GAGGAGCTGGGGTGGCTCGCTCCACCTGCGCAAGCTGCGTAAGCGGCTCCTCCGC 180
QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
Db 181 GATGCCGATGACCTGCAGAAATCCCTGGCAGTGTACAGGCCGGGGCCCGCGAGGGCGGC 240
QY 189 GluArgGlyLeu 192
Db 241 GAGCGGGGCTC 252

Search completed: March 14, 2003, 20:18:35
Job time : 26.2431 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 17:32:58 ; Search time 42.1851 Seconds
(without alignments)
3377.460 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_203
Perfect score: 1014
Sequence: 1 MKVLMALLVTFLGCAKAY.....ARCGAERGLSAIRERLGLV 203

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 501302 seqs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlp
-Q=/cgn2_1/USPNO/spool/US09827854/runtat_11032003_101612_27578/app_query.fasta.1.3576
-DB=Published.Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -MATRIX=blomsum62
-TRANS-human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09827854/ECGN_1_1.298_@runtat_11032003_101612_27578
-NCPU=6 -ICPU=3 -NO_XLPRX -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA: *
1: /cgn2_6/ptodata/1/pubpna/USN7_PUBCOMB.seq: *
2: /cgn2_6/ptodata/1/pubpna/PCJ_NEW_PUB.seq: *
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq: *
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq: *
6: /cgn2_6/ptodata/1/pubpna/PCYUS_PUBCOMB.seq: *
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq: *
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq: *
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq: *
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq: *
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq: *
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq: *
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq: *
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1014	100.0	1156	9	US-09-870-759-129
2	1014	100.0	1156	9	US-09-802-640-17
3	1014	100.0	1156	10	US-09-827-854-8
4	1014	100.0	1291	12	US-10-044-090-454

5	1010	99.6	1156	10	US-09-827-854-12	Sequence 12, Appl
6	1006	99.2	1156	10	US-09-827-854-9	Sequence 9, Appl
7	1006	99.2	1156	10	US-09-827-854-11	Sequence 11, Appl
8	1005	99.1	1157	10	US-09-954-456-760	Sequence 760, Appl
9	1005	99.1	1157	10	US-09-880-107-2244	Sequence 2244, Appl
10	1002	98.8	1156	10	US-09-827-854-7	Sequence 7, Appl
11	999	98.5	1156	10	US-09-827-854-10	Sequence 10, Appl
12	798.5	78.7	41907	10	US-09-967-013-5	Sequence 5, Appl
13	625	61.6	478	10	US-09-964-824A-374	Sequence 374, Appl
14	625	61.6	478	10	US-09-880-107-2491	Sequence 2491, Appl
15	458.5	45.2	449	10	US-09-960-352-4726	Sequence 4726, Appl
16	458	45.2	356	10	US-09-960-352-5420	Sequence 5420, Appl
17	439.5	43.3	442	10	US-09-960-352-9395	Sequence 9395, Appl
18	435	42.9	414	10	US-09-960-352-4237	Sequence 4237, Appl
19	428.5	42.3	416	10	US-09-960-352-8720	Sequence 8720, Appl
20	422.5	41.7	423	10	US-09-960-352-14047	Sequence 14047, Appl
21	405.5	40.0	409	10	US-09-960-352-5148	Sequence 5148, Appl
22	401.5	39.6	416	10	US-09-960-352-9797	Sequence 9797, Appl
23	401	39.5	253	10	US-09-179-5368-130	Sequence 130, Appl
24	387.5	38.2	425	10	US-09-960-352-3497	Sequence 3497, Appl
25	380.5	37.5	390	10	US-09-960-352-1311	Sequence 1311, Appl
26	378.5	37.4	391	10	US-09-960-352-1278	Sequence 1278, Appl
27	379.5	37.4	392	10	US-09-960-352-10599	Sequence 10599, Appl
28	379.5	37.4	401	10	US-09-960-352-4187	Sequence 4187, Appl
29	375.5	37.0	413	10	US-09-960-352-8042	Sequence 8042, Appl
30	374	36.9	252	10	US-09-879-441-24	Sequence 24, Appl
31	374	36.9	252	10	US-09-796-416-24	Sequence 24, Appl
32	369.5	36.4	404	10	US-09-960-352-2325	Sequence 2325, Appl
33	369.5	36.4	405	10	US-09-960-352-14063	Sequence 14063, Appl
34	351.5	34.7	377	10	US-09-960-352-176	Sequence 176, Appl
35	350.5	34.6	353	10	US-09-960-352-4914	Sequence 4914, Appl
36	350.5	34.6	378	10	US-09-960-352-10040	Sequence 10040, Appl
37	350.5	34.6	386	10	US-09-960-352-10361	Sequence 10361, Appl
38	348.5	34.4	388	10	US-09-960-352-11986	Sequence 11986, Appl
39	345.5	34.1	396	10	US-09-960-352-3457	Sequence 3457, Appl
40	344.5	34.0	377	10	US-09-960-352-10540	Sequence 10540, Appl
41	338.5	33.4	432	10	US-09-960-352-8934	Sequence 8934, Appl
42	338	33.3	372	10	US-09-960-352-6552	Sequence 6552, Appl
43	337.5	33.3	376	10	US-09-960-352-5408	Sequence 5408, Appl
44	335.5	33.1	373	10	US-09-960-352-3681	Sequence 3681, Appl
45	334.5	33.0	369	10	US-09-960-352-3494	Sequence 3494, Appl

ALIGNMENTS

RESULT 1
US-09-870-759-129
Sequence 129, Application US/09870759
Patent No. US20020177551A1
GENERAL INFORMATION:
APPLICANT: TERMAN, David S
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
FILE REFERENCE: 870759
CURRENT APPLICATION NUMBER: US/09/870, 759
CURRENT FILING DATE: 2002-01-14
PRIOR APPLICATION NUMBER: US 60/208, 128
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 166
SOFTWARE: PatentIn version 3.1
SEQ ID NO 129
LENGTH: 1156
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (61)..(1014)
OTHER INFORMATION:
US-09-870-759-129

Alignment Scores:
Pred. No.: 3.37e-100
Score: 1014.00
Percent Similarity: 100.00%
Length: 1156
Matches: 203
Conservative: 0

Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-870-759-129 (1-1156)

QY	1	MeysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaVal	20
Db	61	ATGAAGGTTCTGTGGGCTGCTTCTGTGTACATTCTGCGAGAGATGCCAGCCAAAGTG	120
QY	21	GIuGlnAlaValGIuThrGIuProGIuProGIuLeuArgGlnGlnThrGIuTrpGlnSer	40
Db	121	GAGCAAGGGGTGGAGACAGAGACCCGAGCTGGCCAGCAGACCGAGTGGCAGAGC	180
QY	41	GIyGlnArgTrpGIuLeuAlaLeuGlyIaArgPheTrpAspIyrLeuArgTrpValGlnThr	60
Db	181	GGCCAGGCGCTGGAAATCGCACTGGGTCTTTGGGATTAACCTGGCCCTGGGTGCAGACA	240
QY	61	LeuSerGIuGlnValGIuGlnGIuLeuLeuSerSerGIuValThrGlnGIuLeuArgAla	80
Db	241	CTGTCTGAGCAGAGTGCAGAGAGAGACTGTCACTCCAGGTCAACGAACTGAGAGGCG	300
QY	81	LeuMetAspGIuThrMetIysGIuLeuLysAlaTrpIyrSerGIuLeuGlnGIuGlnLeu	100
Db	301	CTGATGACACGACCAATGAAGAGATTGAAGGCCATAATCGGAACCTGGAGAAACAATG	360
QY	101	ThrProValAlaGlnGIuThrArgAlaArgLeuSerIysGIuLeuGlnAlaAlaGlnAla	120
Db	361	ACCCCGGGGGGAGAGACGGGGCCGTGTCAAGAGAGGTGCAGGGGGCGCAGGCC	420
QY	121	ArgLeuGIuAlaAspMetGIuAspValCysGIaArgLeuValGlnTrpArgGIuValAl	140
Db	421	CGCGTGGGGCGGACATGAGAGAGAGTGTGGCGCCCTGTGTGCAGTACCGCGGCGAGTG	480
QY	141	GlnAlaMetLeuGlnGIuSerThrGIuGIuLeuArgValaArgLeuAlaSerHisLeuArg	160
Db	481	CAGCCCATGTGTCGGCCAGAGACCGAGAGAGACTGGGGGTGGCCCTGCGCTCCACCTGCCG	540
QY	161	LysLeuArgIyrAspArgLeuLeuArgAspAlaAspAspLeuGlnIyrAspArgLeuAlaValTrp	180
Db	541	AAGCTGGCTAAGCGGCTCTCTCCGAGATCCGATGACTCTGCAGAAAGGCCCTGGGACGTGTAC	600
QY	181	GlnAlaGlyAlaArgGIuGIuAlaGluArgGIyLeuSerAlaIleArgGluArgLeuGly	200
Db	601	CAGCGCCGGGGCCCGCGAGGGGGCGAGACGGGCGCTTACGCCCATCCGCGAGCGCTGGGG	660
QY	201	ProLeuVal 203	
Db	661	CCCCCTGGGTG 669	

RESULT 2
MS-09-80

US 09 802 040 1/
; Sequence 17, Application US/09802640
; Publication No. US20030036057A1

; GENERAL INFORMATION:

APPLICANT: Bonsai Ariuna

APPLICANT: Kleyon Patrick

TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH

FILE REFERENCE: 24736-2048

CURRENT APPLICATION NUMBER: US/09/802,640

; CURRENT FILING DATE: 2001-03-09

NUMBER OF SEQ ID NOS: 122

```

; SOFTWARE: pas
;
; SEO ID NO 17
;

```

LENGTH: 1156

TYPE: DNA

ORGANISM: *Homo sapiens*

NAME/KEY: CDS

; LOCATION: (61)...(101

OTHER INFORMATION: NU

OTHER INFORMATION: (APOE)
US-09-802-640-17

Alignment Scores:

Pred. No.:	3,37e+100	Length:	1156
Score:	1014.00	Matches:	203
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-802-640-17 (1-1156)

QY	1	MeysValIleuTPAlaIaleuIeuValThrPheIeuIaGlyCysGlnAlaIysVal	20
Db	61	ATGAAGGTTCTGTGGCTGCGCTTCTGTGTCATCTCTGGCAGATCCAGCCMAAGTG	120
QY	21	GIuGlnAlaValGIuThrGIuProGIuProGIuIeuArgGlnGlnThrIuTrpGlnSer	40
Db	121	GAGCAAGGCGTGGAGACAGAACCCGAGCTGGCGCCAGCAGACCGAGTGGCAGAGC	180
QY	41	GIyGlnArgTrpGIuIeuAlaIeuGIyArgPheTrpAspIyIreIeuArgTrpValGIuThr	60
Db	181	GGCCAGCGCTGGGAACGTGGCACTGGTCGCTTTGGGATTAACCTGGCGCTGGGTGCAGACA	240
QY	61	IeuSerGIuGlnValGIuGlnGIuIeuIeuSerSerGlnValThrGlnGIuIeuArgAla	80
Db	241	CTGTCTAGCAGAGTGCAGAGAGAGACTGTCTACGTCCAGGTCAACCCAGAACTGAGGGCG	300
QY	81	IeuMetAspGIuThrMetIySGIuIeuIySalATyIySerGIuIeuGlnGIuGlnIeu	100
Db	301	CTGATGGACGAGACCATGAAGAGATTGAAGGCCATCAAAATCGGAACGTGAGGAACAAC TG	360
QY	101	ThrProValAlaGluGluThrArgAlaArgIeuSerIySGIuIeuGlnAlaIaGlnAla	120
Db	361	ACCCGGGGGGGGAGAGACCCGGGCAKGGGTGTCCAAAGAGCTGCAGGGGGCCAGGCC	420
QY	121	ArgIeuGlnAlaAspMetGIuAspValCysGIyArgIeuValGlnTyArgGIyGlnVal	140
Db	421	CGCGTGGGGCGGACATGGAGAGACTGTGGCGGCGCTGGTGCAGTACCGGCGAGAGTG	480
QY	141	GlnAlaMetIeuGlnIyGlnSerThrGlnGIuIeuArgAlaArgIeuAlaSerHisIleuArg	160
Db	481	CAGGCATGTCTCGGCGCAGAGACCCGAGAGACTGGGGGTGGCTCGCTCCCACTGGCCG	540
QY	161	LysIeuArgIyAsArgIeuIeuArgAspAlaAspAspIeuGlnIyAsArgIeuAlaValTyr	180
Db	541	AAGTGGCGTAAGCGGCTCTCTCCGGATGCCGATGACCTGCAGAAAGCCCTGGCAGTGTAC	600
QY	181	GlnAlaGlyAlaArgGIuGlnIyAlaGlnIyArgGIyIeuSerAlaIleArgGIuArgIeuGly	200
Db	601	CAGCGCCGGGGCCCGCGAGGGCGCCGAGACCGGGCGCTCAGCGCCATCCGCGAGAGCCCTGGGG	660
QY	201	ProIeuVal 203	
Db	661	CCCTGGTG 669	

RESULT 3

US-09-82/-854-8
; Sequence 8, Application US/09827854
; Patent No. US20020123093A1

; GENERAL INFORMATION:

APPLICANT: Zannis, Vassilis

```

; AFFELICANI: nylapicos, nylapicos E.
: TITLE OF INVENTION: Compounds and
:

```

1. TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia

FILE REFERENCE: 07180/004003

CURRENT FILING DATE: 2001-04-05

PRIOR APPLICATION NUMBER: US 09/679,088

PRIOR FILING DATE: 2000-10-04

PRIOR FILING DATE: 2000-04-06

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84


```

QY.      201  Proleval  203
          |||||||||
Db       661  CCCCTGGSTG  669

RESULT 7
US-09-827-854-11
: Sequence 11, Application US/09827854
: Patent No. US20020123093A1

GENERAL INFORMATION:
APPLICANT: Zannis, Vassilis
APPLICANT: Kyriacos, Kyriakos E.
TITLE OF INVENTION: Compounds and methods for lowering
FILE REFERENCE: 07180/004003
CURRENT APPLICATION NUMBER: US/09/827,854
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: US 09/679,088
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 09/544,386
PRIOR FILING DATE: 2000-04-06
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 1156
TYPE: DNA
ORGANISM: Homo sapiens
US-09-827-854-11

```

Pred. No.:	2,44e-99	Length:	1156
Score:	1006.00	Matches:	202
Percent Similarity:	99.51%	Conservative:	0
Best Local Similarity:	99.51%	Mismatches:	1
Query Match:	99.21%	Indels:	0
DB:	10	Gaps:	0

QY	1	MetLysValLeuThrPAlaIalaLeuLeuValThrPheLeuIaIaLysCysGlnIaIaLysVal	20
Db	61	ATGAAGCTTCTGTGGGCTGCGTCTGCTGATCAATCTCTGGCAGAGATCCAGGCCAAGGTG	120
QY	21	GIuGIuIaIaValGIuThrGIuIuProGIuIuLeuArgGlnIuThrGIuTrpGIuSer	40
Db	121	GACCAAGCGGTTGGTAAGACAGACCCGAGCCGAGCTGGCCAGACAGCCAGTGGCGAAGC	180
QY	41	GIyGlnArgTrpGIuLeuIaIaLeuGIuAArgPheThrAspTyrLeuArgTrpAlaIuThr	60
Db	181	GGCCACGCGCTGGGAAGCTGGCACTGGTGGCTTTGGGATTTACCTGGCGCTGGGTGCAGACA	240
QY	61	LeuSerGIuGIuIuValGIuGIuGIuLeuLeuSerGIuIuValIuThrGIuIuLeuArgIa	80
Db	241	CTGTCTGACACAGTGTCCAGAGAGACTCTCAAGTCCAGTCCAGGTCCAGGAACTGAGGGCG	300
QY	81	LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGlnIuLeu	100
Db	301	CTGATGAGAGAAACCATTAAGAGACTTGAAGCGCTACAAATCGGAACTGGAGAACAACTG	360
QY	101	ThrProValAlaGIuGIuThrArgAlaIArgLeuSerLysGIuLeuGlnIaIaGIuIa	120
Db	361	ACCCCGCTGGCGAGAGAACCGGGGCAACGCGCTCTCCAAAGAGACTGCAGCGCGCAGAGCC	420
QY	121	ArgLeuGIuAlaAspMetGIuAspValCysGIuArgLeuValGIuThrArgLysGIuVal	140
Db	421	CGGCTGGGCGCGACATATGAGAGAGCTGTGGCGGCCCTGTGTGTAAGTACGGCGGCAAGTGTG	480
QY	141	GIuIaIaMetLeuGIuGIuIuSerThrGIuGIuIuLeuArgValArgLeuIaIaSerHisLeuArg	160
Db	481	CAGGCGCATGCTCGGCGAAGCAGCAGCAGAGAGACTCGGGGTGCGCTCGCTCCACACTGGCGC	540
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuIaIaValTyr	180
Db	541	AAGCTGTGTGAAGCGCTCTCTCCGATGCGCGATGAGACTCTGAGAAAGCGCTGGCACTGTGAC	600

QY	181	GlnAlaIcylValaaArgGluGlValAGl uArgGlyLeuSerAlaIzIeaArgGluArgGly	200
Db	601	CAGGGCCGGGGCCCGCAGAGGGGCCGACGCCTCAAGCCGCATCCCGCAGCGCTGGG	660
QY	201	ProLeuVal 203	
Db	661	CCCCCTGGGTG 669	

RESULT 8
US-09-954-456-760

Sequence /60, Application US/09954456
Patent No. US20020115057A1
GENERAL INFORMATION:

GENERAL INFORMATION:
APPLICANT: Young, F
TITLE OF INVENTION:

NUMBER OF PATENTS 20000 FOR INVENTIONS WITH UNITED STATES PATENT OFFICE
 TITLE OF INVENTION: Sets
 FILE REFERENCE: 689290-76

; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18

;
; PRIOR APPLICATION NUMBER: US/60/233,617
;
; PRIOR FILING DATE: 2000-09-18

PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: US/60/234,923
 ; PRIOR FILING DATE: 2000-09-25
 PRIOR APPLICATION NUMBER: US/60/235,134

PRIOR APPLICATION NUMBER: US/60/235 637

PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638

PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711

; PRIORITY DATE: 2000-09-27
; PRIORITY APPLICATION NUMBER: US/60/235,720

;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: US/60/235,840

; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,863
 ; PRIOR FILING DATE: 2000-09-27

```

/ FRICTION FILING DATE: 2000 05 21
; NUMBER OF SEQ ID NOS: 2276
: SOFTWARE: PatentIn version 3.0

```

```

; SEQ ID NO 760
; LENGTH: 1157

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens

```

US-09-954-456-760

Alignment scores:
Pred. No.: 3.
Score: 10

Score:	10
Percent Similarity:	99
Best Local Similarity:	99

```

Query match: 99
DB: 10

```

US-09-827-854-15_COPY_1_20

1 MettlysvallEutpa

Db 62 ATGAAGTCTCTGGG

QY 21 GUGNAlaValGluT

Db 122 GAGCAAGCGGTGAGA

QY 41 GLYGLNARGTRPGLUL
|||||

Db 182 GCCACGCGTGGAAC

```

QY      61 LeuSerGluGlnValG
          |||||
hb      243 CmcCmcCmcCmcCmcCmcC

```

UU 242 CUGUCGAGGUGCAGGAGGAGUUGLCACGICCCAAAGTCALCCAAAGAACUGAGGGCG 301

Db 61 ATCAAGGTTCTGTGGGCTGGCTGTGTGTCGTCACATTCCTGGCAGAGATGCCAGGCCAAGCTG 120

Qy 21 GIUGInAlaValGInThrGIuProGInuProGInuLeuArgGInGInuThrGIuTrpGInSer 40

Db 121 GAGCAGACGGTGTGGAGACACAGCCGGAGCCCGAGCTCGCCAGACAGACCGAGTGGCAGAGC 180

Qy 41 GYGInARTrpGInuEnuAlaEnuGlyARpPheTrpAspTrpLeuARTrpValGInThr 60

Db 181 GCCCAGCGCTGGGAATGCGACTGGGTGCTTTGGGATTACCTGCGCTGGGTGCACACA 240

Qy 61 LeuSerGIuGIuValGInGInuGIuLeuLeuSerSerGInValThrGInuGIuLeuArgAla 80

Db 241 CMTGTGACAGAGTGCAGAGAGAGCGTGCCTACGCTCCAGGTACACCCAGAACTGAGGGCG 300

Qy 81 LeuMetASPGLuThrMetLysGInuLeuLysAlaTyrLysSerGIuEnuGIuGInuGIuLeu 100

Db 301 CTGAATGACAGACCATGTAAGAGTGTGAAGAGGCTTACAAATCCGAACTGAGAAACAATCTG 360

Qy 101 ThrProValAlaGInuGIuThrArgAlaArgLeuSerLysGIuEnuGInuAlaGInAla 120

Db 361 ACCCGCGTGGCGAGAGACGCGGGCAGCGCTGTCCAAAGAGCTGACAGGGGGCGACGCC 420

Qy 121 ArgLeuGIuAlaAspMetGInuAspValCysGlyArgLeuValGInuTrpArgGIuVal 140

Db 421 CGGCGTGGCGCGACATGAGAGACGTGCGCGCGCGCTGTGTGACATGCCGGCGAGAGTG 480

Qy 141 GInAlaMetLeuGIuGInSerThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg 160

Db 481 CAGGCCATGCTCGCGCAGACACCGAGGAGCTGCGGGGTGCGCTCCCTCCACTGCGC 540

Qy 161 LysLeuArgLysARLeuLeuArgAspAlaAspAspLeuGInLysArgLeuAlaValTyr 180

Db 541 AAGCTGCGTAACGGCTCTCGCGATGCGAGTGCAGAGACTGTCAGAAAGCGCTGCGCATGTAC 600

Qy 181 GInAlaGlyAlaArgGIuGIuAlaGIuArgGIuLeuSerAlaIleArgGIuArgGIuLeuGIu 200

Db 601 CAGCGCGGGGCGCGGAGGGCGCGAGCGCGGCTCAGCGCCATCCGCGAGCGCTGTGGG 660

Qy 201 ProLeuVal 203

Db 661 CCCCTGGTG 669

RESULT 11

US-09-827-854-10

Sequence 10, Application US/09827854

Patent No. US20020123093A1

GENERAL INFORMATION:

APPLICANT: Zannis, Vassilis

APPLICANT: Kyriacos, Kyriakos E.

TITLE OF INVENTION: Compounds and methods for lowering

TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia

FILE REFERENCE: 07180/004003

CURRENT APPLICATION NUMBER: US/09/827,854

CURRENT FILING DATE: 2001-04-05

PRIOR APPLICATION NUMBER: US 09/679,088

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 09/544,386

PRIOR FILING DATE: 2000-04-06

NUMBER OF SEQ ID NOS: 19

SOFTWARE: FASTSeq for Windows Version 4.0

SEQ ID NO 10

LENGTH: 1156

TYPE: DNA

ORGANISM: Homo sapiens

US-09-827-854-10

Alignment Scores:	
Pred. No.:	1,38e-98
Score:	999.00
Percent Similarity:	99.018
Best Local Similarity:	99.018
Query Match:	98.528
DB:	10
Length:	1155
Matches:	201
Conservative:	2
Mismatch:	0
Indels:	0
Gaps:	0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-827-854-10 (1-1156)

Qy	1	MettysValLeuTPpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaVal	20
Db	61	ATGAAGGTTCTGTGGGGCTGGCTCTGTGCATTCTCGGAGGATGCCAGGCCAAAGTG	120
Qy	21	GlutAlaValAlaGluThrGluProGluProLeuLeuArgGlnGlnThrGluTrpGlnSer	40
Db	121	GAGCAAGCGGTGGAGACAGACAGACCCGAGCTGGCCAGCAGACCACTGGCAAGC	180
Qy	41	GlyGlnArgTrpGluLeuAlaLeuGlyLysPheTrpAspTyrLeuArgTrpValAlaThr	60
Db	181	GGCCAGCGCTGGGAACCTGGCACTGGGTGCTTTTGGATTACCTGGCTGGGTCAACA	240
Qy	61	LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80
Db	241	CTGTCTGTGGCAGGTGCAGGAGAGACTGCTGCAGCTCCACAGTCAACCCAGAACTGAGGGCG	300
Qy	81	LeuMetAspGluTrpMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu	100
Db	301	CTGATGGACGAGACCAATGAAGAGTGAAGGCTTACAAATCGGAACGTGGAGAACTG	360
Qy	101	ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	361	ACCCCGGTGGCGGAGAGAACCCGGGCGACGCTGTTCCAAAGAGCTGCAGCGGCGCCAGGCC	420
Qy	121	ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal	140
Db	421	CGCGTGGGCGCGGACATGAGAGAGCTGTGGGGCGCTGTGGAGTACCGGGCGAGGTG	480
Qy	141	GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	481	CAGCCATGCTCGACAGACAGACCGAGAGAGCTGGGGGTGGCCCTCCACACTGCC	540
Qy	161	LysLeuArgLysAspLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	541	AAGCTGCTTAAGCCGCTCTCTCCGCATGCCGATACCTGCCAAGAGTCCCTGGCAGTGTAC	600
Qy	181	GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgAlaArgLeuGly	200
Db	601	CAGCGCGGGGCGCGGAGGGCGCGAGCGGCGCTCAGAGCGCATCCGAGAGCGGCTGGGG	660
Qy	201	ProLeuVal 203	
Db	661	CCCCCTGTGTG 669	
RESULT 12			
US-09-967-013-5			
: Sequence 5, Application US/09967013			
: Patient No. US20020045840a1			
: GENERAL INFORMATION:			
: APPLICANT: Stanton, Jr, Vincent P.			
: TITLE OF INVENTION: METHOD FOR GENETIC ANALYSIS OF APOE DNA			
: TITLE OF INVENTION: AND USES THEREOF			
: FILE REFERENCE: 11926-022001			
: CURRENT APPLICATION NUMBER: US/09/967,013			
: CURRENT FILING DATE: 2000-10-25			
: PRIOR APPLICATION NUMBER: 60/206,613			
: PRIOR FILING DATE: 2000-05-23			
: NUMBER OF SEQ ID NOS: 91			
: SOFTWARE: FastSeq for Windows Version 4.0			
: SEQ ID NO 5			
: LENGTH: 41907			
: TYPE: DNA			
: ORGANISM: Homo sapiens			
: US-09-967-013-5			

Alignment Scores:	
Pred. No.:	4,63e-75
Score:	798.50
Percent Similarity:	48.19%
Best Local Similarity:	48.19%
Length:	41907
Matches:	186
Conservative:	0
Mismatches:	6

Query Match:	78.75%	Indels:	194
DB:	10	Gaps:	1
US-09-827-854-15_COPY_1_203 (1-203) x US-09-967-013-5 (1-11907)			

OY	12	PheLeuAlaIagIcyGyGclnAlaIalysAlaGluglnAlaValAlGltThgluProGluProGlu	31
Db	20316	TTCCACACAGAGTGCACAGCCCAAGGTGGAGCAAGCGGTGGAGACAGAGCGGAGCCCGG	203795
OY	32	LeuArGInGlnThGltuTrpGlnSerGlyGlnArGtRgLuLeuAlaLeuGlyArGphe	51
Db	20376	CTGGCCGACAGACCGAGTGGAGCGGCGCCAGGGCCTGGGAACCTGGACCTGGGTGGCTTT	204355
OY	52	TrpAspTrpLeuArGtRpyAlaGlnThrIleuSerGluGlnValGlnGluLeuLeuLeuSer	71
Db	20436	TGGATrTACtGtCCCTGGGTGGAGACACTGtCTGAGCAGGtGtCAGAGAGAGtGtCTCAGC	204959
OY	72	SerGlnValAlaThrGlnGluLeu	78
Db	20436	TTCCAGAGTACCCACAGGAACtGAGtGtGAGtGtCCCATCCTGGCCCTTGACCTCTCGGTG	205555
OY	78		78
Db	20556	GACGAGCTrTACCTCCACAGTCCAGGTTTCATTTGCCCCGTGCGCTAAGTCTGGGGG	206155
OY	78		78
Db	20616	CCTGGGCTCTGCTGGTGTCTAGCTTCCTTCCTCCATTCCTGCATTCCTGCTTACCTCTC	206755
OY	78		78
Db	20676	TGGATrTCTCTCTCTCAGACTTGTGTCTCTCTCTCTCCCTCTGAGCTCAGTCTCTACACT	207355
OY	78		78
Db	20736	CGTCCtGGCTCTGtCTCTGtCTGtCCTTCCCTAGCTCTTTTATATAGAGACAGAGATGGGT	207959
OY	78		78
Db	20796	CTCAGCTGTGTGCCAGAGCTGTGTGAACTTCTGGGCTCAAGCGATCTCCCGCCTCGG	208555
OY	78		78
Db	20856	CCTCCCAAGTGTCTGGATTAGAGCAATGAGGACACTTGCCTGGGCTCCTAGCTCTTCT	209155
OY	78		78
Db	20916	TGCTCTGCTCTGCTGCCCCCTTGCACTGTGCTGCACTGTGCTGtCTGtCTCTCTCG	209755
OY	78		78
Db	20976	GCCTTGCCCCGCTTCTTCTCTCTCTCTTGGGTCTCTCTGCTCATTCCCCATCTGCCCCG	210355
OY	78		78
Db	21036	CCCCATCCACGCCCTTCTCCCGGCTCCACATGTGCGACACCCCTCCGCCCTCTCGGCGG	210959
OY	79	ArgAlaIeuMetAspGluThrHemellyGluIeuIysAlaTrpIysSerGluIeuGlnG	98
Db	21096	CAGGGCTGTATGACGACGAGACCATGAAAGAGTTGAAGGCTTACAATGGAATGGAACTGGAA	211555
OY	98	uGlnIeuThrProValAlaGluGluThrArgAlaArgLeuSerIysGluIeuGlnAla	118
Db	21156	ACAACTGACCCCCGCGGGGGGAGAGACCGGGGACAGGCTGTCCAAAGAGCTCGAGGGCGG	212155
OY	118	aglnAlaIArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgG	138
Db	21216	GCAGCCCGGCTGGGCGGCGACATGAGAGAGCTGTGCGGCGCGCTGtGtACATACCGCGG	212755
OY	138	gylIValAlaIleuMetLeuGlyGlnSerThrGlnGluIeuArgValaIArgLeuAlaSerTh	158
Db	21276	CGAGGTGAGGCTCAATGCTGGGCGACAGACCGAGAGAGCTCGGGTGTGCGCTTGCCA	213355

QY	158	stleuNtrqVlyLseuAaGlyLysArlLeuLenuNrgyspAlaAspAspLeuGlnLysArlLeuNl	178
Db	21336	CCTGGCGAAGCTGYGTAAACGGGCTCTCCCGCATGCCATGACCTGACAAAGATGGCTGGC	21395
QY	178	aValIyrrGlnAlaGlyAlaArlGlnLysAlaLysArgGlyLeuSerAlaAlaArlGlnAr	198
Db	21396	ATGTGTAACCAAGGCGGGGGCCCCGGAGGGCGCCAGCGCGGCTCTAGCGCCATCCGCGAGCG	21455
QY	198	gLeuGlyProLeuVal	203
Db	21456	CCTGGGGCCCCCTGGTG	21471

RESULT 13

```

US-09-964-824A-374/c
; Sequence 374, Application US/09964824A
; Patent No. US20020102531A1
;
; GENERAL INFORMATION:
; APPLICANT: Horigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964, 824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236, 033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 374
;
; LENGTH: 478
;
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(478)
; OTHER INFORMATION: n-a,t,g or c
US-09-964-824A-374

```

Alignment scores:	
Pred. No.:	6,38e-59
Score:	625.00
Percent Similarity:	97.69%
Best Local Similarity:	96.92%
Query Match:	61.64%
DB:	10
Length:	478
Matches:	126
Conservative:	1
Mismatches:	2
Indels:	1
Gaps:	0
Gaps:	0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-964-824A-374 (1-478)

Qy	1	MeLysValLeuITPAlaIaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	336	ATGAAGTTCGTGGCGCTCTGGTCAATCTCGGAGGATCCAGGCCAGAGTG	337
Qy	21	GlulAlaValAlgluThrGlulProGlulLeuArgGlnInhrGluTrpGlnSer	40
Db	336	GAGCAAGCGGTGGAGACAGACCGGAGCAGCTGGCCAGCAGACGAGTGGCAGAGC	277
Qy	41	GlyGlnArgTrpGluLeuAlaLeuGlyValArgPheTrpAspTyrLeuArgTrpValGlnThr	60
Db	276	GGCCAGCGCTGGGAAGTGGCACTGGGTGCTTTGGGATTAACCTGCCTGGGTGCACACA	217
Qy	61	LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80
Db	216	CTGTGTGACAGGTGCAGAGAGACCTCTCGAGCTCCACAGTACCCAGAACTGAGAGGC	157
Qy	80	AlaMetAspGluThrMetLeuGlyGluLeuAlaTyrLysSerGluLeuGlnGluLeu	100
Db	156	GCTATGTGACAGACCATGAAGAGTTGAAGCCTACAAATTCGAACTGGAGGACAACT	97
Qy	100	uThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:09:47 ; Search time 825.891 Seconds

(without alignments)
3980.771 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_203

Perfect score: 1014

Sequence: 1 MKVLMALVTFVLAGCCQAKV.....AREGAEGLSAIRERLGPLV 203

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Command line parameters:
-MODE=frame+pn.model -DEV=xlp
-Q/cgn2.1/USPTO/US0982785/Runat_11032003_101610_27495/app.query.fasta_1.3576
-DB=EST -QFWT=fastap -SUFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNIT5=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=DCT -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0982785.acgn.1.1.8826_runat_11032003_101610_27495 -NCPU=6 -ICPU=3
-NO_XLPPX -NO_MMAP -LARGEUSER -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_din:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	%	Match	Length	DB	ID	Description
1	1014	100.0	690	13	B1599793	B1599793	603245228	B1599793	603245228
2	1014	100.0	699	12	B6769455	B6769455	602742580	B6769455	602742580
3	1014	100.0	748	12	B6706913	B6706913	602672086	B6706913	602672086
4	1014	100.0	748	13	B1553085	B1553085	603196058	B1553085	603196058
5	1014	100.0	757	13	B1603658	B1603658	603251982	B1603658	603251982
6	1014	100.0	757	13	BM042228	BM042228	603616186	BM042228	603616186
7	1014	100.0	790	13	B1551811	B1551811	603192716	B1551811	603192716
8	1014	100.0	798	12	B6708414	B6708414	602672374	B6708414	602672374
9	1014	100.0	800	13	BM042094	BM042094	603615713	BM042094	603615713
10	1014	100.0	817	12	B6774871	B6774871	602649975	B6774871	602649975
11	1014	100.0	845	12	BG829472	BG829472	602763768	BG829472	602763768
12	1014	100.0	933	13	B1668318	B1668318	603295681	B1668318	603295681
13	1014	100.0	965	14	B0677266	B0677266	603295681	B0677266	603295681
14	1011	99.7	692	13	B1546740	B1546740	603191304	B1546740	603191304
15	1010	99.6	802	13	B1458355	B1458355	603198705	B1458355	603198705
16	1010	99.6	811	13	B1600563	B1600563	603244936	B1600563	603244936
17	1010	99.6	942	12	B1600906	B1600906	603249241	B1600906	603249241
18	1009	99.5	797	12	BG715366	BG715366	602675503	BG715366	602675503
19	1009	99.5	808	13	B1668329	B1668329	603295692	B1668329	603295692
20	1007	99.3	741	12	B6762924	B6762924	602735153	B6762924	602735153
21	1006	99.2	725	12	BG708285	BG708285	602675940	BG708285	602675940
22	1006	99.2	748	12	BG715607	BG715607	602675940	BG715607	602675940
23	1006	99.2	803	13	B1670350	B1670350	603292721	B1670350	603292721
24	1006	99.2	842	12	BG763371	BG763371	602735433	BG763371	602735433
25	1006	99.2	891	13	B1548292	B1548292	603189930	B1548292	603189930
26	1006	99.2	1027	13	B1670367	B1670367	603292738	B1670367	603292738
27	1005	99.1	790	12	BG707147	BG707147	602670283	BG707147	602670283
28	1005	99.1	938	12	BG761746	BG761746	602717942	BG761746	602717942
29	1004	99.0	677	13	B1604042	B1604042	603243857	B1604042	603243857
30	1004	99.0	790	13	B1551066	B1551066	603196425	B1551066	603196425
31	1002	98.8	1100	14	BM914382	BM914382	AGENCOCURT	BM914382	AGENCOCURT
32	1001	98.7	687	12	BG706243	BG706243	602669540	BG706243	602669540
33	1001	98.7	693	12	BG707355	BG707355	602672757	BG707355	602672757
34	1001	98.7	812	13	B1601551	B1601551	603250760	B1601551	603250760
35	1000	98.6	856	13	BM042032	BM042032	603616028	BM042032	603616028
36	999	98.5	1030	14	B0682454	B0682454	602675874	B0682454	602675874
37	998	98.4	705	12	BG715556	BG715556	602675874	BG715556	602675874
38	998	98.4	919	13	B1551475	B1551475	603194314	B1551475	603194314
39	997	98.3	907	12	BG706129	BG706129	602669093	BG706129	602669093
40	993	97.9	941	14	B0681329	B0681329	AGENCOCURT	B0681329	AGENCOCURT
41	992	97.8	680	10	AV702686	AV702686	AV702686	AV702686	AV702686
42	989	97.5	1023	14	B0680425	B0680425	AGENCOCURT	B0680425	AGENCOCURT
43	988	97.4	669	10	AV655989	AV655989	AV655989	AV655989	AV655989
44	987	97.3	716	12	BG703146	BG703146	602686035	BG703146	602686035
45	987	97.3	757	12	BG703504	BG703504	602686279	BG703504	602686279

ALIGNMENTS

RESULT 1
LOCUS B1599793
DEFINITION 603245228F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5287710 5',
mRNA sequence.
ACCESSION B1599793
VERSION B1599793
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

Db 57 ATGAAGTTCTGTGGCTGCTGGTGCATCTCTGGCAGAGATGCCAGGCAAGGTG 116
Qy 21 GtuglnalavalglutThrGlupProgluleuArgrnglnThrGlutPglInser 40
Db 117 GAGCAAGGGGTGAGACAGACGCCGAGCTGCCCGACGACGAGCACTGGCAGAGC 176
Qy 41 GlylnArgrtPgluleuAlaleuGlyArpPheTrpaspTyrLeuArgrtPvalGlnThr 60
Db 177 GGCAGCGCTGGAACTGGACCTGGGTGCTTTGGGATCTACCTGGCTGGGTGAGACA 236
Qy 61 LeuSerGluglnValGlnGlnGlnLeuSerSerGlnValThrGlnGlnLeuArGAla 80
Db 237 CTGCTGGAGCAGGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 296
Qy 81 LeuMetaspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
Db 297 CTGATGGACGAGACCATTAAGAGAGTGAAGGCTTCAAAATGGAACTGGAACTG 356
Qy 101 ThrProValAlaGlnGlnThrArgAlaArgrLysSerLysGlnLeuGlnAlaGlnAla 120
Db 357 ACCCGGGTGGGAG 416
Qy 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgrLeuValGlnTyrArgGlyGlnVal 140
Db 417 CGGCTGGGCGCGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 476
Qy 141 GlnAlaMetLeuGlnGlnInserThrGlnGlnLeuArGValArGLeuAlaSerHisLeuArG 160
Db 477 CAGGCGATGCTCGGCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 536
Qy 161 LysLeuArGlyArgrLeuLeuAlaAspAlaAspAspLeuGlnLysArGLeuAlaValTyr 180
Db 537 AAGCTGGCTAAGCCGCTCTCCGCGATGCCATGACCTGCAGAAAGCGCTGGCAGTGTAC 596
Qy 181 GlnAlaGlyAlaArgrGlnGlnAlaGlnArGlyLeuSerAlaIleArGlnArGLeuGly 200
Db 597 CAGGCGGCGGCGCGAGGCGCGAGCGCGCTCAGCGCCATCCCGAGCGCGCTGGGG 656
Qy 201 ProLeuVal 203
Db 657 CCCCTGGTG 665

RESULT 3
LOCUS BG706913 748 bp mRNA linear EST 07-MAY-2001
DEFINITION 602672086F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4794789 5',
mRNA sequence.
ACCESSION BG706913
VERSION BG706913.1 GI:13982733
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 748)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@femail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshitaki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM0676 row: 1 column: 22
High quality sequence stop: 748.
Location/Qualifiers
1. 748

FEATURES
source
/organism="Homo sapiens"

/db_xref="taxon:9606"
/clone="IMAGE:4794789"
/clone_1db="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: Brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dt primed using primer 5'-tttttttttttttttttttt-3',
size-selected for average insert size 2.3 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 136 a 224 c 289 g 99 t
ORIGIN

Alignment Scores:
Pred. No.: 2e-105 Length: 748
Score: 1014.00 Matches: 203
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x BG706913 (1-748)

Qy 1 MetLysValLeuThrPAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 73 ATGAAGTTCTGTGGCTGCTGGTGCATCTCTGGCAGAGATGCCAGGCAAGGTG 132
Qy 21 GtuglnalavalglutThrGlupProgluleuArgrnglnThrGlutPglInser 40
Db 133 GAGCAAGGGGTGAGACAGACGCCGAGCTGCCCGACGACGAGCACTGGCAGAGC 192
Qy 41 GlylnArgrtPgluleuAlaleuGlyArpPheTrpaspTyrLeuArgrtPvalGlnThr 60
Db 193 GGCAGCGCTGGAACTGGACCTGGGTGCTTTGGGATCTACCTGGCTGGGTGAGACA 252
Qy 61 LeuSerGluglnValGlnGlnGlnLeuSerSerGlnValThrGlnGlnLeuArGAla 80
Db 253 CTGCTGGAGCAGGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 312
Qy 81 LeuMetaspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
Db 313 CTGATGGACGAGACCATTAAGAGAGTGAAGGCTTCAAAATGGAACTGGAGAGAACTG 372
Qy 101 ThrProValAlaGlnGlnThrArgAlaArgrLysSerLysGlnLeuGlnAlaGlnAla 120
Db 373 ACCCGGGTGGGAG 432
Qy 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgrLeuValGlnTyrArgGlyGlnVal 140
Db 433 CGGCTGGGCGCGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 492
Qy 141 GlnAlaMetLeuGlnGlnInserThrGlnGlnLeuArGValArGLeuAlaSerHisLeuArG 160
Db 493 CAGGCGATGCTCGGCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 552
Qy 161 LysLeuArGlyArgrLeuLeuAlaAspAlaAspAspLeuGlnLysArGLeuAlaValTyr 180
Db 553 AAGCTGGCTAAGCCGCTCTCCGCGATGCCATGACCTGCAGAAAGCGCTGGCAGTGTAC 612
Qy 181 GlnAlaGlyAlaArgrGlnGlnAlaGlnArGlyLeuSerAlaIleArGlnArGLeuGly 200
Db 613 CAGGCGGCGGCGCGAGGCGCGAGCGCGCTCAGCGCCATCCCGAGCGCGCTGGGG 672
Qy 201 ProLeuVal 203
Db 673 CCCCTGGTG 681

RESULT 4
B1553085

LOCUS B1553085 748 bp mRNA linear EST 05-SEP-2001
 DEFINITION 603196058p1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5275358 5',
 mRNA sequence.
 ACCESSION B1553085
 VERSION B1553085.1 GI:15440397
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 748)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 AUTHORS Contact: Robert Strausberg, Ph.D.
 EMAIL: cga@bbs-r@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM11695 row: d column: 15
 High quality sequence stop: 748.
 Location/Qualifiers
 1..748
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5275358"
 /clone_lib="NIH_MGC_95"
 /tissue_type="hippocampus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescriptR (modified
 pBluescript KS+); Site.1: BamHI, Site.2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
 size-selected for average insert size 2.5 kb and
 normalized to R0T 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 136 a 227 c 288 g 97 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2e-105 Length: 748
 Score: 1014.00 Matches: 203
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0
 US-09-827-854-15_COPY_1_203 (1-203) x B1553085 (1-748)

QY 1 MettysvalleutrpAlaAlaLeuValThrpheleuAlaGlyCysGlnAlaVal 20
 |||||
 DB 75 ATGAAGGTTCTGTGGCTGCTGTGCATTCCTGCGAGGATGCCAGGCCAAGG 134
 |||||
 QY 21 GtlnAlaAlaValGlturThrcInuProGltuLeuAtrGlnGlnThrcInuTrpGlnSer 40
 |||||
 DB 135 GAGCAAGCGGTGAGACAGAGCCGCGAGCTGCCGACAGCCAGTGGCAGAGC 194
 |||||
 QY 41 GtGlnATGTrpGltuLeuAlaLeuGlyArgpHetTrpAspTrpLeuAtrGtrpValGlnThr 60
 |||||
 DB 195 GGCACGCGGTGGAACTGGCACTGGGTGCTTTGGGATTACCTGGCTGGGCGCAGACA 254
 |||||
 QY 61 LeuSerGtlnAlaValGtlnGtlnuLeuLeuSerSerGlnValThrcInuLeuAtrGlnAla 80
 |||||
 DB 255 CTGCTGAGACAGTGCAGAGAGAGCTGCACCTCCACAGTCAACCAGGAAGTGAAGGGG 314
 |||||
 QY 81 LeuMetAspGlturThrcMetCysGltuLeuLysAlaTrpLysSerGtlnuLeuGtlnGtlnuLeu 100

DB 315 CTGATGAGACAGACCATGAGAGAGGCTCAATCGAACTGAGGAGCAACATCG 374
 |||||
 QY 101 ThrProValAlaGltuThrAtrGlnAlaArgLeuSerTrpGltuLeuGlnAlaAla 120
 |||||
 DB 375 ACCCGGTGGCGAGAGAGACGCGGCGACGCTGTCTCAAGGACTCGACGCGCGCAGGGC 434
 |||||
 QY 121 ArgLeuGtlnAlaAspMetGltuAspValCysGtlnArgLeuValGlnTrpArgGtlnVal 140
 |||||
 DB 435 CGGCTGGCGCGACATGAGAGACGCTGGCGCCCTGTGCTACATACCGCGCAGAGTG 494
 |||||
 QY 141 GtlnAlaMetLeuGtlnSerThrcInuLeuAtrGtlnArgLeuAlaSerTrpLeuAtr 160
 |||||
 DB 495 CAGGCGATGCTGGCGCAGACACCGAGAGCTGGCGGTGCTGCTCCACCTGGCGC 554
 |||||
 QY 161 LysLeuAtrGtlnArgLeuAtrGtlnAlaAspAspLeuGtlnLysArgLeuAlaValTrp 180
 |||||
 DB 555 AAGCTGCGTAACGCGCTCTCCGCGATGCCGATGACCTCAACAGCGCTGGCACTGTAC 614
 |||||
 QY 181 GtlnAlaGtlnAlaAtrGtlnGtlnAlaGtlnArgGtlnLeuSerAlaTrpArgGtlnuLeuGtln 200
 |||||
 DB 615 CAGGCGGAGGCGCGGAGGCGCGCGGCGCTCAGCGCCATCCGCGAGCGCTGGGG 674
 |||||
 QY 201 ProLeuVal 203
 |||||
 DB 675 CCCCTGGTG 683

RESULT 5
 B1603658
 LOCUS B1603658 757 bp mRNA linear EST 07-SEP-2001
 DEFINITION 603251982p1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5303396 5',
 mRNA sequence.
 ACCESSION B1603658
 VERSION B1603658.1 GI:15496597
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 757)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 AUTHORS Contact: Robert Strausberg, Ph.D.
 EMAIL: cga@bbs-r@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM11768 row: d column: 21
 High quality sequence stop: 753.
 Location/Qualifiers
 1..757
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5303396"
 /clone_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescriptR (modified
 pBluescript KS+); Site.1: BamHI, Site.2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
 size-selected for average insert size 2.3 kb and
 normalized to R0T 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 138 a 230 c 292 g 97 t

DB 530 AAGCTGCTAAGCGGCTCCTCCGCGATGCCGATGACCTGCAGAAAGCGCTGCAGTGTAC 589

QY 181 GlnAlaGlyAlaArgGluGluGlyAlaArgGlyLeuSerAlaIleArgGluArgGluGly 200
|||||
DB 590 CAGCGCGGCGCCGCGAGCGCGCGCTCAGCGCATCCGCGAGCGCTGGG 649

QY 201 ProLeuVal 203
|||||
DB 650 CCCCTGGTG 658

RESULT 7
BI551811 790 bp mRNA linear EST 05-SEP-2001
LOCUS 603192716F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:526415 5',
DEFINITION mRNA sequence.
ACCESSION BI551811
VERSION BI551811.1 GI:15439123
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 790)
NIH-MGC http://mhc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM1665 row: P column: 04
High quality sequence stop: 777.
Location/Qualifiers
1. 790
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:526415"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: Brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
size-selected for average insert size 2.5 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH-MGC Library."

BASE COUNT 141 a 235 c 311 g 103 t

ORIGIN

Alignment Scores:
Pred. No.: 2.16e-105 Length: 790
Score: 1014.00 Matches: 203
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x BI551811 (1-790)

QY 1 MellyValleuTrpAlaAlaLeuValThPhenLeuAlaGlyCysGlnAlaVal 20
|||||
DB 75 ATGAAGTTCTGTGGCTGGTGTGTCACATTCCTGCGAGATGCCAGGCCAAGTGT 134

QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40

DB 135 GAGCAAGCGGTGAGACAGAGCGCGAGCCCGAGCTCGCCAGCAGACCGAGTGGAGAC 194
|||||

QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||

DB 195 GGCACCGCTGGGAACTGGACATGGGGTCCCTTTTGGATTACTCGGCTGGGTGAGACA 254
|||||

QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||

DB 255 CTGTCTGAGAGAGTGCAGAGAGAGCTGCTCAGCTCCAGGTCCAGAGAACTGAGGGG 314
|||||

QY 81 LeuMetAspGluThrMetLeuGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
|||||

DB 315 CTGATGAGACAGACCATGAAGAGTGAAGCCCTACAAATCCGAACTGGAGGAACTAG 374
|||||

QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerGluLeuGlnAlaAlaGlnAla 120
|||||

DB 375 ACCCGGTGGCGGAGAGAGCGGGGACGGCTGTCCAGAGAGCTGCAGGGCGCGAGGCC 434
|||||

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||

DB 435 CGGCTGGCGCGGACATGAGAGACGTGTGGCGCGCTGTGCAGTACCGCGCGAGGTG 494
|||||

QY 141 GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||

DB 495 CAGGCCATGCTCGCGCAGAGACCGAGAGCTCGGGTGGCTCCCTCCACCTCGCGC 554
|||||

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||

DB 555 AACCTGCGTGAAGCGGCTCCTCCCGCATGCCGATGACTGCAGAAAGCGCTGCAGTAC 614
|||||

QY 181 GlnAlaGlyAlaArgGluGluGlyAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||

DB 615 CAGCGCGGCGCGCGAGGCGCGCGAGCGCGCTCAGCGCCATCCGCGAGCGCTGGG 674
|||||

QY 201 ProLeuVal 203
|||||

DB 675 CCCCTGGTG 683

RESULT 8
BG708414 798 bp mRNA linear EST 07-MAY-2001
LOCUS 602672374F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4794955 5',
DEFINITION mRNA sequence.
ACCESSION BG708414
VERSION BG708414.1 GI:13985733
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 798)
NIH-MGC http://mhc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM10677 row: C column: 20
High quality sequence stop: 778.
Location/Qualifiers
1. 798
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4794955"
/clone_lib="NIH_MGC_96"

Oy	121	ArgLeuGlyAlaAspMetGluAspValGlySerGlyArgLeuValGlnTyrArgGlyIleVal	140
Db	410	CGGGTGCGCCGCAGATGAGAGCAGTGTGCCGCCCTGTGTCACTACCGCGGCGAGGTG	469
Oy	141	GlnAlaMetLeuGLGIserThrGluGluLeuAArgValArgLeuAlaserHisLeuArg	160
Db	470	CAGGCCATGCTCTGGCCAGAGCACCGAGAGACTGCGGGGTGCGCTTCCTCCCACTGCGCC	529
Oy	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspleuGlnLysArgLeuAlaValTyr	180
Db	530	AAGCTGCGGAAGCGGCTCTCTCCGCATGCGCATGTACCTGCACAACGCCCTGGCAGTTAC	589
Oy	181	GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGlnArgLeuGly	200
Db	590	CAGGCGGGGCCCCCGAGAGGCGCCGAGGCGGCGCTCAGCGCATCCGCGAGCGCCTGGGG	649
Oy	201	ProLeuVal	203
Db	650	CCCCGTGGTG	658
RESULT	10		
LOCUS	Bg774871	817 bp	mRNA linear EST 15-MAY-2001
DEFINITION	602649975P1 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:4760827 5'		
ACCESSION	Bg774871		
VERSION	Bg774871.1	GI:14045188	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 817)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cga@bsf-research.nih.gov Tissue Procurement: DCM/DMP CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LNCMI612 row: e column: 20 High quality sequence stop: 813.		

FEATURES	SOURCE	LOCATION/QUALIFIERS
BASE COUNT		148 a 242 c 324 g 103 t
ORIGIN		Note: This is a NIH-MGC Library."
		1. .817
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone="IMAGE:4760827"
		/clone_lib="NIH-MGC:40"
		/tissue_type="carcinoma, cell line"
		/lab_host="DH10B (phage-resistant)"
		/note="Organ: prostate; Vector: pORF7; Site_1: XhoI;
		Site_2: EcoRI; cDNA made by oligo-dT priming.
		Directionally cloned into EcoRI/XhoI sites using the
		following 5' adaptor: GGCACGAG(G). Library constructed by
		Ling Hong in the laboratory of Gerald M. Rubin (University
		of California, Berkeley) using XP-cDNA synthesis kit
		(Stratagene) and Superscript II RT (Life Technologies).

Alignment Scores:	
Pred. No.:	2.26e-105
Score:	1014.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Length:	817
Matches:	203
Conservative:	0
Mismatches:	0

Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0
US-09-827-854-15_COPY_1_203 (1-203) x BG774871 (1-817)			

OY	1	MethylsValIleuTriPAlaIaleuIeuValThrPheIeuAlaG1CYsGlnIalAlaYsVal	20
Db	40	ATGAAGGTTCTGTGGGCTCCGTTGCTGGTACATTCTCGCAGATGCCAGGCCAAAGTG	99
OY	21	GIuGlnAlaIaValaGIuThrcIuProG1uPurG1uIeuAArgG1nG1nThrcIuTrpG1nSer	40
Db	100	GAGCAAGCGGTGGAGACAGAGAGCCGAGGCCGAGCTGGCCACAGACCAAGCGCAGAGC	159
OY	41	GIyGlnIaArgTrpG1uIeuAlaIeuGlnIaArgPheTrpAspTrpIeuAArgTrpAlaGlnThr	60
Db	160	GCCCAAGCGCTGGGAAGCTGGCACTGGGTGGCTTTGGATTACTTCGCTGGGTGGACAA	219
OY	61	LeuSerGIuGlnIaValaG1nG1uG1uIeuIeuSerSerG1nIaValThrcIuG1uIeuAArgAla	80
Db	220	CTGTCTGAGACAGCTGCAGAGAGAGCTGCTGCACCTCCAGGTCAACCAGCAATGAGGGCG	279
OY	81	LeuMetAspG1uThrMetIySG1uIeuYsAlaTrpYIySerGIuIeuG1nG1uIeuIeu	100
Db	280	CTGATGGACAGAACCATGAGAGATTGAAGGCTTCAATTCGAATCGAATCGAGAGAACACG	339
OY	101	ThrProValaIaG1uG1uThrArgIaIaArgIeuSerIySG1uIeuG1nAlaIaGlnAla	120
Db	340	ACCCCGGTGGCCGAGAGACGCGGGCCACGGCTGTCCAAAGAGCTGCAGCGGGCCAGGCC	399
OY	121	ArgIeuGlnIaIaAspMetG1uAspValCYsG1aArgIeuAlaG1nTrpArgG1yG1uVal	140
Db	400	CGGCTGGGCGCGGACATGAGAGACTGTGTCCGCCGCCCTGTCAGTACCGCGGCAAGTG	459
OY	141	GlnAlaMetIeuG1yG1nSerThrcIuG1uIeuAArgValaArgIeuAlaSerHisIeuAArg	160
Db	460	CAGGCCATGCTCTGCGCAGAGCACCCAGAGAGCTGCGGGGTGCGCTCGCTCCACCTGGCC	519
OY	161	LysIeuAArgIyAspArgIeuIeuAArgAspAlaAspAspIeuGlnIyAspArgIeuAlaValTrp	180
Db	520	AAGCTCGCTAAGCGGCTCTCTCGCATGCGCATGACCTGCACAAAGCGCTGGCACTGTAC	579
OY	181	GlnAlaG1aIaIaArgG1uG1uIaG1uArgIuArgG1yIeuSerAlaIaIaArgG1uArgIeuG1y	200
Db	580	CAGGCGGGGGGCCGCGAGAGGCGCGCAGCGGCGCTCAGCGCCATCCGCGAGCGCTGGGG	639
OY	201	ProIeuVal 203	
Db	640	CCCCGTGGTG 648	

RESULT	11
BG829472	
LOCUS	845 bp mRNA linear EST 22-MAY-2001
DEFINITION	602763768b1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4899112 5'
ACCESSION	BG829472
VERSION	BG829472
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 845)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

CDNA Library Preparation: Ling Hong/Rubin laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.lnl.gov
Plate: L1CM1790 row: 9 column: 17
High quality sequence stop: 829.
Location/Qualifiers

FEATURES

SOURCE

1. 845
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:489112"
/clone_lib="NIH_MGC_42"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site: 1: XhoI;
Site: 2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"

BASE COUNT 153 a 244 c 342 g 106 t
ORIGIN

Alignment Scores:

Align. No.: 2.37e-105 Length: 845
Score: 1014.00 Matches: 203
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Dbs: 12 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x BG829472 (1-845)

OY 1 MettysValleuTPRAlaaleuenuValThrPheleuAlaGlyCysGlnAlaVal 20
Db 57 ATGAAGGTTCTGTGGCGTGGCTGTGTCACATTCCTGGCGAGATGCCAGGCGAAGGTG 116
OY 21 GtuglnAlaValaGluThrGluProGluLeuAraGlnGlnGlnThrGluTrpGlnSer 40
Db 117 GAGCAAGGCGGTGAGACAGAGCGGAGCCGAGCTGCCAGACACAGCGAGTGGCAGAGC 176
OY 41 GlyGlnAraGTrpGluLeuAlaLeuGlyAraGpHetPaspTyrLeuAraGTrpValGlnThr 60
Db 177 GCGCCAGCGCTGGAACTGGCACTGGCTGCTTTGGGATTACCTGCGTGGGTGCAGACA 236
OY 61 LeuSerGluGlnValGlnGluGluLeuSerSerGlnValThrGlnGluLeuAraGln 80
Db 237 CTGTCTGAGCAGGTGCAGAGAGGCTGCTCAGCTCCCAAGGTACCCAGGAACGTGAGGCGC 296
OY 81 LeuMetAspGluThrMetCysGluLeuLysAlaTyrLysSerGluLeuGluGluGln 100
Db 297 CTGATGAGAGACCATATGAGAGGTGAAGGCTTCAAAATCGAAGCTGAGAGAACTG 356
OY 101 ThrProValAlaGluGluThrAraGlnLeuSerLysGluLeuGlnAlaVala 120
Db 357 ACCCGCGGTGGCGAGAGACGCGGCGCTGTCCAAGAGAGCTCAAGCGCGCGCAGAGCC 416
OY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyAraGluLeuValGlnTyrAraGlyVal 140
Db 417 CGGCTGGGCGGAGACATGAGAGCTGTGCGCGCTGTGTCAGTACCGCGCGAGGTG 476
OY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuAraGlnValArgLeuAlaSerHisLeuAra 160
Db 477 CAGGCGATGCTGGCGCAGAGCACCGAGAGCTGGGCTGCGCTCCCACTGCGC 536
OY 161 LysLeuAraGlyAraGluLeuAraGlnAlaAspAspLeuGlnLysAraGlnAlaValaTyr 180
Db 537 AAGCTGCGTAAGCGGCTCTCCGCGATGCTCGAGAAAGCGCTGGCACTGTAC 596
OY 181 GlnAlaGlnAlaAraGluGlnValaGlnAraGlyLeuSerAlaHisLeuGlnAraGln 200
Db 597 CAGGCGGCGGCGCGAGAGGCGCGAGCGCGCTCAAGCGCGCTGCGCGTGGG 656

OY 201 ProLeuVal 203
Db 657 CCCCTGTGTG 665

RESULT 12

B1668318 933 bp mRNA linear EST 12-SEP-2001
LOCUS B1668318
DEFINITION 603295681P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5314844 5',
mRNA sequence.

ACCESSION B1668318
VERSION B1668318.1 GI:15582551
KEYWORDS EST.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 933)
NIH-MGC <http://mgc.nci.nih.gov/>.

AUTHORS

Toshiyuki and Piero Carninci (RIKEN)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

JOURNAL

Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov

COMMENT

Tissue Procurement: Mixios Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
<http://image.lnl.gov>
Plate: L1AM11798 row: a column: 21
High quality sequence stop: 780.

FEATURES

SOURCE

1. 933
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5314844"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site: 1: BamHI; Site: 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 171 a 284 c 359 g 119 t
ORIGIN

Alignment Scores:

Align. No.: 2.73e-105 Length: 933
Score: 1014.00 Matches: 203
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Dbs: 13 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x B1668318 (1-933)

OY 1 MettysValleuTPRAlaaleuenuValThrPheleuAlaGlyCysGlnAlaVal 20
Db 75 ATGAAGGTTCTGTGGCGTGGCTGTGTCACATTCCTGGCGAGATGCCAGGCGAAGGTG 134
OY 21 GtuglnAlaValaGluThrGluProGluLeuAraGlnGlnGlnThrGluTrpGlnSer 40
Db 135 GAGCAAGGCGGTGAGACAGAGCGGAGCCGAGCTGCCAGACACAGCGAGTGGCAGAGC 194
OY 41 GlyGlnAraGTrpGluLeuAlaLeuGlyAraGpHetPaspTyrLeuAraGTrpValGlnThr 60
Db 195 GCGCAGCGCTGGAACTGGCACTGGCTGCTTTGGGATTACCTGCGTGGGTGCAGACA 254

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LRAM1662 row: g column: 01
 High quality sequence stop: 692.
 Location/Qualifiers

FEATURES

Source

1. 692
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5262744"
 /clone_lib="NIH_MGC_95"
 /tissue_type="hippocampus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescriptR (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgaag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.5 kb and
 normalized to 400 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
 BASE COUNT 131 a 207 c 262 g 92 t
 ORIGIN

Alignment Scores:

Pred. No.: 3.94e-105 Length: 692
 Score: 1011.00 Matches: 202
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.51% Mismatches: 0
 Query Match: 99.70% Indels: 0
 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x B1546740 (1-692)

QY 1 MettysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 75 ATGAAGCTTCTGCTGGCTGCGTTCATTCCTGCGCAGATGCCAGCCAAAGCTG 134
 QY 21 GtGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 Db 135 GAGCAAGCGGGTGGAGACAGACCGGAGCGCCAGCGCCAGACAGCAAGTGGCAGAGC 194
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 Db 195 GGGCAGCGCTGGAACTGGCACTGGCTTTGGGATTAACCTGCGCTGGTCAGACCA 254
 QY 61 LeuSerGlnGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 Db 255 CTGCTGAGCAGGTGCAGAGAGAGCTCAGCTCCAGGTCCACAGCAAGTGGAGCGC 314
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 Db 315 CTGATGGAGAGACCATGAAGAGCTTGAAGCTTCAATGCAATCGAAGTGGAGCAACTG 374
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 Db 375 ACCCGGGTGGGAGAGACAGCGGGCAGCGCTGTCCAGAGAGCTCAGCGCGAGGCC 434
 QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlnArgLeuValGlnTyrArgGlyGluVal 140
 Db 435 CGGCTGGCGCGGACATGAGAGAGAGCTGTGCGCGCCCTGTGTCAGTACCGCGCGAGTGTG 494

QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 495 CAGCGCATGCTCGGCGACAGACAGCAGAGCTGGCGGTGCCCTCGCTCCACCTGGC 554
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 Db 555 AAGCTGCTAGACGGCTCTCCGCGATGCCATGCCTCAGAGAGCGCTGGCAGTGTAC 614
 QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaAlaLeuArgGluGlnGly 200
 Db 615 CAGCGCGGGCGCGGAGGCGCGAGCGCGCTCAGCGCCATCCGCGAGCGCTGGG 674
 QY 201 ProLeuVal 203
 Db 675 CCCCTGGTG 683
 RESULT 15
 B1458355 802 bp mRNA linear EST 21-AUG-2001
 LOCUS 603198705F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5278177 5',
 DEFINITION mRNA sequence.
 ACCESSION B1458355
 VERSION B1458355.1 GI:15248998
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 802)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 TISSUE Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LRAM1702 row: j column: 02
 High quality sequence stop: 779.
 Location/Qualifiers

FEATURES

source

1. 802

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5278177"

/clone_lib="NIH_MGC_96"

/tissue_type="hypothalamus"

/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescriptR (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgaag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.3 kb and
 normalized to 400 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT

143 a 238 c 319 g 102 t
 ORIGIN

Alignment Scores:

Pred. No.: 6.31e-105 Length: 802
 Score: 1010.00 Matches: 202
 Percent Similarity: 99.51% Conservative: 0
 Best Local Similarity: 99.51% Mismatches: 1
 Query Match: 99.61% Indels: 0
 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x B1458355 (1-802)


```
OY      1 MetLysValLeuTrpAlaIleuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
      |||||||
Db      75 ATGAAGTTCTGTGGGCTGTGGTGTGTCACATTCCTGGCAGGATGCCAGGCCAAGGTG 134
OY      21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnThrGluTrpGlnSer 40
      |||||||
Db      135 GAGCAAGCGGTGGAGCAGAGCGGAGCCGAGCTGGCCAGCAGACCGAGTGGCAGAGC 194
OY      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
      |||||||
Db      195 GGCACAGCGCTGGGAACCTGGGACTGGGTCTTTGGGATTACTGGCGTGGTGCAGACA 254
OY      61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
      |||||||
Db      255 CTGTCTGACGACAGGTGCAGAGAGCTGTCCAGCTCCAGGTCCAGGAACAGAGGCGC 314
OY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
      |||||||
Db      315 CTGATGACGAGACCATGAAGAGTTGAAGGCTTACAAATCGGAACCTGGAGAACACTG 374
OY      101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
      |||||||
Db      375 ACCCGGTTGGCGGAGAGACGCGGCGACGGCTGTCCAAGAGCTGCAGGGCGCGCAGGCC 434
OY      121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
      |||||||
Db      435 CGGCTGGCGCGGACATGAGAGACGTGTGCGCGCCGCTGGTCAGTACC GCGCGGAGGTG 494
OY      141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
      |||||||
Db      495 CAGGCCATGCTCGGCCAGAGACACCGAGAGACTGCGGGTGGCGCTCGCTCCACCTGGCGC 554
OY      161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
      |||||||
Db      555 AAGCTGCGTAAGGGGCTCTCCCGCATGGCGATGACCTGCAGAGGCGCTGGCAGGTGTAC 614
OY      181 GlnAlaGlyAlaArgGluGluAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
      |||||||
Db      615 CAGGCCGGGGGCGCGAGGGCGCGGAGCGGCGCTCAGCGCCATCCGCGAGCGCGCTGGGG 674
OY      201 ProLeuVal 203
      |||||||
Db      675 CCCCTGGTG 683
```

Search completed: March 14, 2003, 20:13:59
Job time : 829.891 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:08:57 ; Search time 2016.1 Seconds
(without alignments)
3998.551 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_277
Perfect score: 1374
Sequence: 1 MKVLMAALVFLAGCOAKV.....RAKLEBOAQOIRLQAEAFQA 277

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=rlp
-Q=/cgn2_1/USPTO.spool/US09827854.r/unat_11032003_101610_27486/app.query.fasta_1.3576
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNTS-bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USPR=US09827854.ecgn_1.1.13008_@unat_11032003_101610_27486 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOP -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FANGOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: GenEmbl:*
2: gb_ba:*
3: gb_hlg:*
4: gb_in:*
5: gb_om:*
6: gb_ov:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ot:*
22: em_ov:*
23: em_pal:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pin:*
35: em_hlg_rtd:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_hlgo_hum:*
40: em_hlgo_mus:*
41: em_hlgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1374	100.0	1110	6 E00359	E00359 CDNA coding
2	1374	100.0	1110	6 E00823	E00823 DNA sequence
3	1374	100.0	1147	6 AX302545	AX302545 Sequence
4	1374	100.0	1156	6 BD004278	BD004278 Apo E hum
5	1374	100.0	1156	6 BD004278	BD004278 Apo E hum
6	1374	100.0	1186	6 BC003557	BC003557 Homo sapi
7	1367	99.5	1110	6 E08423	E08423 DNA coding
8	1365	99.4	1156	6 BD004277	BD004277 Apo E hum
9	1365	99.3	1157	6 AX333278	AX333278 Sequence
10	1365	99.3	1157	6 AX409597	AX409597 Sequence
11	1365	99.3	1157	6 H15975	H15975 Sequence 1
12	1365	99.3	1157	6 H0MAPOE	M12529 Human apoli
13	1362	99.1	1156	6 BD004279	BD004279 Apo E hum
14	1255	91.3	1178	6 MFAPOE	X13887 Monkey mRNA
15	1185.5	86.3	5497	6 AF261279	AF261279 Homo sapi
16	1185.5	86.3	41907	6 AX358722	AX358722 Sequence
17	1185.5	86.3	41907	6 AF050154	AF050154 Homo sapi
18	1185.5	86.3	107567	9 AC011481	AC011481 Homo sapi
19	1185.5	86.3	5515	9 H0MAPOE4	M10065 Human apoli
20	1142	83.1	5413	9 AF261280	AF261280 Pan trogl
21	1079	78.5	1138	4 AF303830	AF303830 Tupaia gl
22	1067	77.7	208239	2 AC021988	AC021988 Homo sapi
23	1065	77.5	4762	9 BABAPOE	M29522 Baboon apoli
24	1021.5	74.3	1060	4 RABAPOLP	M36603 Rabbit apoli
25	1015	73.9	1045	10 MUSAPOE	M12414 Mouse apoli
26	1015	73.9	1104	10 BC028816	BC028816 Mus muscu
27	1004.5	73.1	965	6 AX384545	AX384545 Sequence
28	1004.5	73.1	1108	4 BPAPOE	X61171 B. taurus mr
29	1004.5	73.1	5617	6 AX384541	AX384541 Sequence
30	1004.5	73.1	6026	6 AX384539	AX384539 Sequence
31	1002.5	73.0	1126	6 AR164342	AR164342 Sequence
32	1002.5	73.0	1126	6 AR205885	AR205885 Sequence
33	996.5	72.5	1122	4 SSAPOE	X72835 S. scrofa mr
34	986.5	71.8	1154	4 BPAPOE	X64839 B. taurus mr
35	984	71.6	660	6 A62340	A62340 Sequence 28
36	984	71.6	660	6 A62342	A62342 Sequence 30
37	984	71.6	660	6 A62344	A62344 Sequence 32
38	984	69.9	951	6 AR075563	AR075563 Sequence
39	960	69.9	951	10 S76779	S76779 rAPOE-apoli
40	943	68.6	718	9 AF20049753	AF20049753 Pan trogl
41	939	68.3	718	9 AF20050053	AF200502 Gorilla g
42	938	68.3	959	10 M0SAPOE	M73490 Mus musculu
43	936	68.1	718	9 AF20050353	AF200505 Pongo pyg
44	932	67.8	718	9 AF20050653	AF200508 Hylobates
45	922.5	67.1	1069	10 RATAPOE	J00705 rat apolipo

RESULT 1

ALIGNMENTS

E00359 LOCUS E00359 1110 bp RNA linear PAT 29-SEP-1997
 DEFINITION cDNA coding human apolipoprotein E3.
 ACCESSION E00359
 VERSION E00359.1 GI:2168646
 KEYWORDS JP 1985118189-A/1.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1110)
 Teranishi, Y., Takamatsu, N., Matsui, Y., Kimura, M. and Ikeda, Y.
 REFERENCE DNA FRAGMENT
 PATENT: JP 1985118189-A 1 25-JUN-1985;
 MITSUBISHI CHEM IND LTD
 COMMENT OS human
 PN JP 1985118189-A/1
 PD 25-JUN-1985
 PF 29-NOV-1983 JP 1983224980
 PI TERANISHI YUTAKA, TAKAMATSU NOBUHIKO, MATSUI YASUSHI, PI
 KIMURA MASAKO,
 PI IKEDA YASUKO
 PC C12N15/00, C07H21/04//C12P21/00;
 CC strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC *source: tissue; location=livr;
 FH Key location/Qualifiers
 FT CDS 15..968
 FT /product='apolipoprotein E3 precursor' FT
 FT sig_peptide 15..68
 FT /product='apolipoprotein E3 signal peptide' FT
 FT mat_peptide 69..965
 FT /product='apolipoprotein E3'.
 FEATURES
 source location/Qualifiers
 1..1110
 /organism='Homo sapiens'
 /db_xref='taxon:9606'
 BASE COUNT 198 a 353 c 416 g 143 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.17e-72 Length: 1110
 Score: 1374.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-827-854-15_COPY_1_277 (1-277) x E00359 (1-1110)
 QY 1 MetlysalLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db |||||||
 QY 15 ATGAAGTTCTGTGGCTGCTCTGTCACATTCCTGCAGATGCCAGCCAAAGTG 74
 Db |||||||
 QY 21 GluGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 Db |||||||
 QY 75 GAGCAACGGGTGGAGACAGACCCGACCTGCGCCGACGACCGAGTGGCAAGC 134
 Db |||||||
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 Db |||||||
 QY 135 GGCAGCGCTGGGAACTGGCACTGCGTTCGCTTTGGGATTACCTGCCCTGGGTGCACACA 194
 Db |||||||
 QY 61 LeuSerLeuGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 Db |||||||
 QY 195 CTGCTCTGAGCAGGTGCAGGAGGACCTCTCACTCCAGGTCCAGGAACTGAGGCG 254
 Db |||||||
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
 Db |||||||
 QY 255 CTGATGACGACGACCATGAAGGAGCTGAAGGCTTCAAAATCGGAATCGAGAACTGAGAACTG 314
 Db |||||||

QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
 Db |||||||
 QY 315 ACCCGCTGGCGGAGGAGAGCGGGCCAGCGCTGCCAAGAGGTGCAGGGCGGAGGCC 374
 Db |||||||
 QY 121 ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
 Db |||||||
 QY 375 CGGCTGGGCGGACATGAGAGACGTGTGGCGGCTGTGTGACGTACCCCGGAGAGTG 434
 Db |||||||
 QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db |||||||
 QY 435 CAGCCATGCTCGGCCAGAGACACGAGAGAGCTGGGGGTGGCGCTCCGCTCCACCTGCG 494
 Db |||||||
 QY 161 LysLeuArgLysArgLeuLeuArgAlaAspAlaAspAspLeuLysArgLeuAlaValTyr 180
 Db |||||||
 QY 495 AAGCTGGTAAAGCGGCTCCCGGATGCCGATGACCTGAGAAAGCCCTGGCAAGTGTAC 554
 Db |||||||
 QY 181 GlnAlaGluAlaArgGluGluGluAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db |||||||
 QY 555 CAGGCGGGGCGCCGAGGGCGCCGAGCGGCGCTCAGCCCATTCGCGAGCGCTGGGG 614
 Db |||||||
 QY 201 ProLeuValGluGlnGluArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
 Db |||||||
 QY 615 CCCCTGTGGAAACAGGGCGCGTGGCGGCGCCACTGTGGGTCCCTCGCGCGGCGAGCGG 674
 Db |||||||
 QY 221 LeuGlnGluArgAlaGlnAlaTTPGlyGluArgLeuArgAlaArgMetGluMetGly 240
 Db |||||||
 QY 675 CTACAGGAGCGGGCGCCAGCGCTGGGGGAGCGGCTGGCGCGGCGGAGTGGAGATGGGC 734
 Db |||||||
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 Db |||||||
 QY 735 AGCCGAGCCCGGACCGCTCGAGAGAGTGAAGAGAGAGTGGCGGAGTGGCGGCAAG 794
 Db |||||||
 QY 261 LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAla 277
 Db |||||||
 QY 795 CTGAGAGAGCAGGCGCCAGCATACGCTGAGCGGCGGCTGCGCGGAGTGGAGGCC 845
 Db |||||||
 RESULT 2
 E00823 1110 bp DNA linear PAT 29-SEP-1997
 LOCUS E00823
 DEFINITION DNA sequence coding for human apolipoprotein E and its signal peptide.
 ACCESSION E00823
 VERSION E00823.1 GI:2169084
 KEYWORDS JP 1986096997-A/1.
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1110)
 AUTHORS Teranishi, Y., Matsui, Y., Ikeda, Y. and Kimura, M.
 TITLE PRODUCTION OF HUMAN APOLIPOPROTEIN E-LIKE PROTEIN
 JOURNAL Patent: JP 1986096997-A 1 15-MAY-1986;
 MITSUBISHI CHEM IND LTD
 COMMENT OS Human (Homo sapiens)
 PN JP 1986096997-A/1
 PD 15-MAY-1986
 PF 16-OCT-1984 JP 1984216987
 PI TERANISHI YUTAKA, MATSUI YASUSHI, IKEDA YASUKO, KIMURA MASAKO
 PC C12P21/00, A61K35/74, A61K37/04, C12N15/00, (C12P21/00, C12R1:19),
 PC (C12N15/00,
 CC strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC *source: tissue; location=livr;
 FH Key location/Qualifiers
 FT 3'UTR 1..14
 FT sig_peptide 15..68
 FT /product='human apolipoprotein E signal peptide' FT
 FT CDS 69..968
 FT /product='human apolipoprotein E' FT 3'UTR

FEATURES	969	1110	Location/Qualifiers
source	1	1110	/organism="unidentified" /db_xref="taxon:32644"
BASE COUNT	198 a	353 c	416 g 143 t
ORIGIN			
Alignment Scores:			
Pred. No.:	2.17e-72	Length:	1110
Score:	1374.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-09-827-854-15_COPY_1_277 (1-277) x E00823 (1-1110)			
OY	1	MettysValleuTrrPaalaaLeuLeuValThrrPheLeuaIagIcysGlnaIatysVal	20
Db	15	ATGAAGTTCTGTGGCGCTTCTGTGCATTCCTGGCAGATCCAGGCCAAGTG	74
OY	21	GIuGlnaIaValaGIuThrrGluProGluLeuAargIcngInThrrGlnSer	40
Db	75	GAGCAAGCGGTGGAGACAGAACCCGAGCCGAGCTCCGACAGACAGAGCGAGTGAGAGC	134
OY	41	GLYGIaAgtTrrPgluLeuAaLeuGlyAaRphetrPasprrPrrLeuAgtTrrPValaGInThr	60
Db	135	GCCACGCCCTGGAACTGGCACTGGGTGCTTTGGGATTAACCTGGCCTGGGTGCAGACA	194
OY	61	LeuSerGIuGlnaIaGInGluGluLeuLeuSerSerGlnaIaThrrGlnIuLeuAargIa	80
Db	195	CTGTCTGACGAGGTGGACAGAGAGACTGTGTCAGTCCAGGTACACCCAGAACTAGAGGGCG	254
OY	81	LeuMetaspGIuThrrMetIysGIuLeuLysaIaTrrIysserGIuLeuGlnGluGlnLeu	100
Db	235	CTGAAGCAGAGACCAATCAAGAGATTGAAGCCCTAACAAATGGAAACGGAGAAACAATG	314
OY	101	ThrrProValaIaGlnGluGluThrrAargIaAargLeuSerIySGIuLeuGlnaIaIaGlnaIa	120
Db	315	ACCCGGGTGGGGAGAGAGACCGGGGACAGCGCTGTCCAAAGAGCTGCAGCGGGCGACGCC	374
OY	121	AargLeuGlyaIaaspMetGIuAspIyAcySGIaArgLeuValaGlnTrrAargIyGluVal	140
Db	375	CGGGTGGGGCGGACATGAAGAGAGTGTGGCGCGCTGTGGTAGATACCAGCGGAGGTG	434
OY	141	GlnaIaMetLeuGlnGlnSerThrrGlnGluLeuAargIaAargLeuAaSerHisIaAarg	160
Db	435	CAGGCAATGCTCGGCCAGACAGACCCGAGAGAGTGGGGGTGGCTCGCTCCACCTGGCC	494
OY	161	LysLeuAargIyAargLeuLeuAargAspAaAspAspLeuGlnIyAargLeuAaValaTrr	180
Db	495	AAGCGCTTAAGCGGCTCTCCGGAATCGCATCACTGCAGAAAGCCCTGCAGAGTGTAC	554
OY	181	GlnaIaGlyaIaAargGlnGluGlnaIaGlnaIaAargIyLeuSerHisIaIaAargIyGluVal	200
Db	555	CAGGCCGGGGGGCGGAGAGGGCGGAGAGGGGGCTTCAGGCCCAATCCGCGAGCCCTGGGG	614
OY	201	ProLeuValaGlnGlnGlyAargIaValaIaAaIaThrrAaGlySerLeuAaIaGlnPro	220
Db	615	CCCCGTGGTGAACAGGGGGCGGTGGGGCGGCCCACTGTGGGTCTCCCTGGCGGCGAGCG	674
OY	221	LeuGlnGluAargIaGlnaIaIaTrrGlyGluAargLeuAargIaAargMetGluMetGly	240
Db	675	CTACAGGAGCGGGCCAGAGCCCTGGGGGAGACGGCTGGCGCGCGGAGAGGAGATGGCG	734
OY	241	SerAgtTrrAargAspAargLeuAspGluValaIySGlnGlnIaValaIaGluValaAargIaIyS	260
Db	735	AGCGGAGACCGCGAACCGCTTGAGAGAGTGAAGAGACAGAGTGGCGGAGAGTGGCGCCAG	794
OY	261	LeuGlnGluGlnaIaGlnGlnIaIaIyLeuGlnaIaGlnaIaAaPheGlnaIa	277
Db	795	CTTGAGAGGACAGGCCAGACAGATACCGCTTGAGGCGCGAGGCTTTCACAGGCC	845

RESULT 3																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

QY	201	ProleValaGlucInglYrValaArqAlaIatrrAlaIglYserLeuaIaGlYInPro	220
Db	646	CCCCTGTGTGAACAGAGCGCGCTGGCGGCGCCACTGTGTGGCTCCCTGGCGGCGACCG	705
QY	221	LeuGIngluArqAlaGlnAlaIatrrpGlYluArqLeuaArqAlaIarMetGlunIuMetGly	240
Db	706	CTACAGACCGGGCCAGCGCTGGGGCGAGCGGCTGCCGATGGAGAGATGGGC	765
QY	241	SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArqAlaLys	260
Db	766	AACCGGACCGCGGACCGCTCGACAGAGGAAAGAGACAGGTGGCCGAGGTGGCCCAAG	825
QY	261	LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla	277
Db	826	CTGGAGGACGAGCCCGACGATACGCTCGAGCCGAGGCTTCCAGGCC	876
RESULT 4			
LOCUS	BD004278	1156 bp	DNA Linear
DEFINITION	Apo E humanized mammal.		
ACCESSION	BD004278		
VERSION	BD004278.1	GI:1863239	
KEYWORDS	JP 2001017028-A/2.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1156)		
TITLE	Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.		
JOURNAL	Apo E humanized mammal Patent: JP 2001017028-A 2 23-JAN-2001;		
COMMENT	MITSUBISHI CHEMICAL CORP		
	OS Homo sapiens (human)		
	PN JP 2001017028-A/2		
	PD 23-JAN-2001		
	PF 28-APR-2000 JP 2000128919		
	PR SHINOBU FUJITA,HIROKI HAWANAKA,YUKO FUKUI,MINESUKE YOKOYAMA		
	PI A01K67/027,A51K45/00,A61P25/28,A61P43/00,C12N5/10,PC		
	C12N15/09//C07K14/775,		
	PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,(C12N5/00,C12R1:91) CC		
FEATURES			
source	FT Key Location/Qualifiers (61)..(1011).		
	1..1156 Location/Qualifiers		
	/organism="Homo sapiens" /db_xref="taxon:9606"		
BASE COUNT	208 a 368 c 432 g 148 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	2.26e-72	Length:	1156
Score:	1374.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-09-827-854-15_COPY_1-277 (1-277) x BD004278 (1-1156)			
QY	1	MeltyValIleuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	61	ATGAAGGTTCTGTGGCGCGCTGTGTGCATTCCTCGGCGAGATGGCAGCCAGGTG	120
QY	21	GlunGlnAlaValaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer	40
Db	121	GAGCAAGCGGTGAGACAGAGCCGCGAGCTGCGCCAGACAGACCGAGTGGCAGAGC	180
QY	41	GlyGlnArgTrpGluLeuAlaLeuGluLysArgPheThrPaspArgTrpLeuArgTrpValGlnThr	60
Db	181	GCCAGCCCTGGGAAGCTGCACTGGTGGCTTTTGGATTAACCTGCGCTGGGTGCAGACA	240

QY	61	LeuSerGluGlnValGInGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80
Db	241	CTGTCTGACACAGGTGCAGAGAGGCTGCTCACTGCTCCAGGATCACCCAGAACTGAGGGCG	300
QY	81	LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu	100
Db	301	CGATGTGACAGACCATGAAAGAGTTGAAGGCGCTACAAATTCGAACTGAGGAACTG	360
QY	101	ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	361	ACCCCGGTGGCGAGAGAGACGGGGCACGCTGCTCCAAAGAGACTCAGGCGGCGAGGCC	420
QY	121	ArgLeuGluAlaAspMetLysAspValLysGlyArgLeuValGlnTyrArgGlyGluVal	140
Db	421	CGGCTGGGGCGGACACTGGAGAGACTGTGTGGCCCGCTGTGTCACTACCGCGGAGGTG	480
QY	141	GlnAlaMetLeuGlnLysInSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	481	CAGGCCATCTCGGCCACAGACACCGAGAGAGTGCAGGGTGCCTGCGCTCCACACTGGCG	540
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	541	AAGCTCCGTAACCGGCTCTCCGCGATGCCGATGACCTGCACAAACCGCTGGCACTGTAC	600
QY	181	GlnAlaGlyAlaArgGluGlnLysAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200
Db	601	CAGCGCGGGGCGCCGAGAGGGCGCCGACGCGGCTCAGCGCCATCCGCGAGGCGCTGGGG	660
QY	201	ProLeuValGluGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro	220
Db	661	CCCCTGTGTGACAGAGGCGCGCTGCGGGCGCGCACATGTGCGCTCCCTGGCGGACAGCG	720
QY	221	IleuGlnGlnThrAlaGlnAlaThrProLysGluArgLeuArgAlaArgMetGluIleuMetGly	240
Db	721	CTACAGGAGACGGGGCCAGGCGCTGGGGCGACGCGGCGCGCGGATGAGAGATGGGCG	780
QY	241	SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAlaArgAlaLys	260
Db	781	ACCCGAGACCGCGGACCGCTGTGACAGCAGGTGAAGAGCAGGTGGCGAGGTGCGGCCAAG	840
QY	261	IleuGlnGlnGlnAlaGlnGlnIleArgLeuGlnGlnAlaGlnAlaIleGlnAla	277
Db	841	CTGGAGGAGACAGGCCAGACAGTACGCTGCGAGGCGGAGCGCTTCACAGGCC	891
RESULT 5			
HUMAPOE3		1156 bp	linear PRI 24-NOV-2000
LOCUS			
DEFINITION	Homo sapiens preapolipoprotein E (APOE) mRNA, complete cds.		
ACCESSION	K00396		
VERSION	K00396.1 GI:178850		
KEYWORDS	apolipoprotein; apolipoprotein E; lipoprotein; polymorphism; very low density lipoprotein.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 355 to 1156) Breslow,J.L., McPherson,J., Nussbaum,A.L., Williams,H.W., Loftquist,Kahl,F., Karathanasis,S.K. and Zannis,V.I.		
TITLE	Identification and DNA sequence of a human apolipoprotein E cDNA clone.		
JOURNAL	J. Biol. Chem. 257 (24), 14639-14641 (1982)		
MEDLINE	83082756		
PUBMED	6897404		
REFERENCE	2 (bases 250 to 777)		
AUTHORS	Wallis,S.C., Rojane,S., Gill,L., Markham,A., Edge,M., Woods,D., Williamson,R. and Humphries,S.		
TITLE	The isolation of cDNA clones for human apolipoprotein E and the detection of apoE RNA in hepatic and extra-hepatic tissues		
JOURNAL	EMBO J. 2 (12), 2369-2373 (1983)		
MEDLINE	84131952		
PUBMED	6199196		

REFERENCE AUTHORS	3 (bases 1 to 1156) Zannis-V.I., McPherson,J., Goldberger,G., Karathanasis,S.K. and Breslow,J.L.	CDS	61..1014 /gene="APOE" /codon_start=1 /product="preapolipoprotein E" /protein_id="AA859546.1" /db_xref="GI:178851" /db_xref="GDB:G00-119-691" /translation="MKYLMAALVTFLAQCQAKVEQAVETEPPELPKQTEWMSGORM ELAGRFMDYLRWYOTLSEVOEELSSQVQELRALMDETMKELKAYKSELOLP VAETFRARLSKELOAAQARIGADMEDYCGRLVOYRGVQAMIGOSTEELRYRLASHLR KLKRLRLDADDDOKRLAVYQACAREGARELSLTIRELGTPLVQGRVRAVTGSLAG QPLDERAOWEERLRAREWGSRTFRDLDEKRVQVAERAKLEQAOQIRLQLEAFO ARLSWEPEPLVEDQROWAGLVEKVOAAVGTSAAPVPSDNH" 61..114 /gene="APOE" /note="apolipoprotein E" 115..1011 /gene="APOE" /product="apolipoprotein E mature peptide" 267 /gene="APOE" /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3], [4]" /citation=[2] /replace="c" 279 /gene="APOE" /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3]" /citation=[4] /replace="a" 288 /gene="APOE" /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3]" /citation=[4] /replace="a" 369 /gene="APOE" /note="in variant" /citation=[4] /replace="a" 409 /gene="APOE" /note="in variant" /citation=[4] /replace="a" 448 /gene="APOE" /note="in type III hyperlipoproteinemia" 538 /gene="APOE" /note="in type III hyperlipoproteinemia" 568 /gene="APOE" /note="in variant" /citation=[4] /replace="c" 586 /gene="APOE" /note="in epsilon-2 allele" /citation=[3] /citation=[5] /replace="t" 702 /gene="APOE" /note="c in [J. Biol. Chem. 258, 11422-11422 (1983)], [3], [4]" /citation=[2] /replace="g" 783 /gene="APOE" /note="in variant"
TITLE	Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing of a new variant	sig_peptide	
JOURNAL MEDLINE	J. Biol. Chem. 259 (10), 6498-6504 (1984)	mat_peptide	
REFERENCE AUTHORS	5 (bases 577 to 624) Gill,L.L., Peoples,O.P., Pearson,D.H., Robertson,F.W., Humphries,S.E., Cumming,A.M. and Hardman,N.	variation	
TITLE	Isolation and characterisation of a variant allele of the gene for human apolipoprotein E		
JOURNAL MEDLINE	Biochem. Biophys. Res. Commun. 130 (3), 1261-1266 (1985)		
REFERENCE AUTHORS	6 (sites) Rall,S.C. Jr., Newhouse,Y.M., Clarke,H.R., Weisgraber,K.H., McCarthy,B.J., Mahley,R.W. and Bersot,T.P.	variation	
TITLE	Type III hyperlipoproteinemia associated with apolipoprotein E phenotype E3/3. Structure and genetics of an apolipoprotein E3 variant		
JOURNAL MEDLINE	J. Clin. Invest. 83 (4), 1095-1101 (1989)	variation	
COMMENT	2539388 [1] corrected in [J. Biol. Chem. 258, 11422-11422 (1983)]. [J. Biol. Chem. 258, 11422-11422 (1983)] correction of [1]. [4] [5] epsilon-2 allele. [6] sites: mutations resulting in type III hyperlipoproteinemia. Apo E is a component of normal human very low density lipoprotein. There are six human apo E phenotypes known to result from a single structural gene, three of the common alleles being epsilon-4, epsilon-3 and epsilon-2. This sequence appears to be of the epsilon-3 allele. [1] argues that the apo E polymorphism involves mutations in the structural coding region; for example the epsilon-2 phenotype which is characterized by hyperlipoproteinemia is thought to result from a c to t change (arg to cys) at base 586 below [3]. [5]. The sequence shown is 578 homologous with human apo A-I and 818 homologous with rat apo E. For the epsilon-4 sequence, see the separate entry. [J. Biol. Chem. 258, 11422-11422 (1983)] is too extensively revised relative to [1] to record the revisions in the FEATURES table. The sequence below is that of [J. Biol. Chem. 258, 11422-11422 (1983)] and [3]. Apo E is located on chromosome 19 --Jackson, Bruns and Breslow, PNAS USA 81, 2945-2949 (1984) -- and is believed to be linked to the Apo C-II gene (see separate entry). The two mutations causing type III hyperlipoproteinemia (apolipoprotein E phenotype E3/3) produces substitutions of Arg for Cys at amino acid 112 and Cys for Arg at amino acid position 142. Draft entry and printed copy of sequence for [1] kindly provided by L.L. Gill, University of Aberdeen. Complete source information: Human liver [1] [J. Biol. Chem. 258, 11422-11422 (1983)], [2], [3], [4] and blood [5], cDNA to mRNA. Location/Qualifiers 1..1156 /organism="Homo sapiens" /db_xref="taxon:9606" /map="19q13.2" /tissue_type="liver and blood" 1..1156 /gene="APOE"	variation	
FEATURES	source	variation	
gene		variation	

Db	495	AAAGCTGCGTAAAGCGGCTCTCTCCGCGATGCGCATGACCTGCGAAGCGGCTGGCAATGTAC	554
Qy	181	GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200
Db	555	CAGCGCGGGGCCCCGAGGGCGCGCAGCGCGGCTCAGCGCGCATCGCGAGGCGCTGGGG	614
Qy	201	ProLeuValGluGluGlyArgValAlaArgAlaAlaThrValGlySerLeuAlaGlyGlnPro	220
Db	615	CCCCGTGTGGAAACAGGGCGCGGTGCGGGCCGCCATGTGTGGCTCCCTGTGGCCGAGCGG	674
Qy	221	LeuGlnGluArgAlaGlnAlaArgPylGluIuArgLeuArgAlaArgMetGluGluMetGly	240
Db	675	CTACGAGAGCGGGGCCCAAGGCTGGGGGCCAAGCGGCTGGCCGCGCGGATGGAGAGATGGGC	734
Qy	241	SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys	260
Db	735	AGCCGAGACCCGCGACCGCTGGACGAGGTGAAGGAGCAGGTGGCCGAGGTGGCGGCCAAG	794
Qy	261	LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAla	277
Db	795	CTGGGAGGAGCAGGGCCACAGATACGCTCGGAGGCGCGGCTTCCAAGGC	845

RESULT	8				
LOCUS	BD004277				
DEFINITION	BD004277	1156 bp	DNA	Linear	PAT 31-JAN-2002
ACCESSION	BD004277				
VERSION	BD004277.1	GI:18632238			
KEYWORDS	JP 2001017028-A/1.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 1156)	Mammalia, Eutheria, Primates; Catarrhini; Homnidae; Homo.			
	Fujita, S., Hamanaka, H., Fukui, Y. and Yokoyama, M.			
	Apo E humanized mammal			
	Patent: JP 2001017028-A 1 23-JAN-2001;			
	MITSUBISHI CHEMICAL CORP			
	OS Homo sapiens (human)			

EN 25-JAN-2001
ED 23-JAN-2001
PF 28-APR-2000 PR 2000128919
PI
SHINOBU FUJITA, HIROKI HAWANAKA, YUKO FUKU, MINESUKE YOKOYAMA PC
A01667/027, A61K45/00, A61P25/28, A61P43/00, C12N5/10, PC
C12N15/09, //CO7K14/775,
PC C12N15/00, C12R1:91), C12N5/00, C12N15/00, (C12N5/00, C12R1:91) CC

FEATURES	source	Location/Qualifiers	Location/Qualifiers
		1. .1156	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
BASE COUNT		208 a	367 c 432 g 149 t
ORIGIN			

Alignment Scores:		
pred. No.:	6.65e-72	length: 1156
Score:	1366.00	Matches: 276
Percent Similarity:	99.64%	Conservative: 1
Best Local Similarity:	99.64%	Mismatches: 0
Query Match:	99.42%	Indels: 0
DB:	6	Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x BD004277 (1-1156)

QY 1 MetIysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20

Db 61 ATGAGGTTCTGTGGGCTGCGTTGCTGTCACATTCTGCGAGGATGCCAGCCCAAGTG 120

21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40

Db 121 GAGCAGCGGTGAGACAGACCGAGCCGAGCTGCCTCAGCAGACCCAGTGCCAGACC 180

QY 41 GlyGlnArgTrrpGluLeuAlaLeuGlyArgPheTrrpAspTrrLeuArgTrrpValGlnThr 60

Db 181 GGCACGCGTGGAACTGGCAGTGGTCGCTTTTGGATTACCTGCCCTGGGTGCAGACA 24

```
61 LeUserSergLucInValGInglucInLeuLeuSersSergInValThrInglucInLeuArgAla 80
```

241 C1G1C1G0ACAGG1G0CAAGAGAGC1G0C1CAAGC1CCCAAG1CAGCCCAAGAGAGC1G0AGAGAG 301

01 Define and point in the city separately
02
03
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99

```
101  mbrDecval[1][aC][uG][uThb+rva]aArOfouSortVec[uLauG]na)a)aG]na)a 122
```

361 ACCCGGTTGGCGGAGGACGCGCGGCACTGTTTCCAAAGGAGTTGACGGCGGCGCAGGCC 422

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 14

Db 421 CGGCTGGGGCGGACATGGAGGACGTGTGCGGCCCGCTGTGCAGTACCGCGGCGAGGTG 48

QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 166

Db 481 CAGGCATGCTCGGCCAGAGACCGAGGAGCTGGGGTGGGCTCGCCTCCACCTGGCC 54

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValI Tyr 188

D6 541 AACCTGCCGTAGCGGCTCCTCCGGGATGCCGATGACCTGCAGAGTGCCTGCCAGTGTAC 60

[illegible][illegible][illegible]

321 TauC[nc]uArra}aG[na]aTrng[uc]uArriuArra}aAroMet[uc]uMet[v]24

Dh 721 CTACAGAGAGCGGGGGCCAGGGCTGGGGCGAGCGGCTCGCGCGCGGATGAGAGGATGGCC 78

QY 241 SerarqThrarqasparqLeuaspqIuVallysgIugInValAlaqluValarqAlalys 26

Db 781 AGCCGACCCGCGACCGCTTGACGAGGTGAAGGAGCAGGTGGCGGAGGTGCGGCCCAAG 84

QY 261 LeuGlulGlulnAlaGlulnIleArgLeuGlulnAlaGlulnAla 277

Db 841 CTGAGGAGCAGGCCACAGATACGCCCTGCAGGCCGAGGCCCTTCCAGGCC 891

RESULT 9
AX333278

LOCUS	AX3332/8	115/ bp	DNA	linear	PAI 09-JAN
DEFINITION	Sequence 3787 from Patent WO0194629.				

ACCESSION	AA333276
VERSION	AX333278.1
	GI:18123912

SOURCE human.

Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Eutheria; Primates; Catarrhini; Hominoidea; Homo

REFERENCE
1
Young P. E., Augustus M., Carter K. C., Ebner R., Endress G.

TITLE Cancer gene determination and therapeutic screening using siq

gene sets
JOURNAL Patent: WO 0194629-A 3787 13-DEC-2001;

FEATURES	AVAILON Pharmaceuticals (US)
Location/Qualifiers	

```

SOURCE
1: 1157
/organism="Homo sapiens"

```


BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN

Alignment Scores:
Pred. No.: 7 62e-72 Length: 1157
Score: 1365.00 Matches: 275
Percent Similarity: 99.28% Conservative: 0
Best Local Similarity: 99.28% Mismatches: 2
Query Match: 99.34% Indels: 0
DB: 6 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x AX333278 (1-1157)

QY 1 MettysValleuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20
DB 62 ATGAAGGTTCTGTGGGCTGCTGTGGTGCATCTCTGCGAGAGTCCAGGCCAAGGTG 121

QY 21 GluGlnAlaValaGluThrGluProGluLeuValGlnThrGluTrpGlnSer 40
DB 122 GAGCAAGCGGTGAGACAGAGCGGAGCCGAGCTGCGCAGACAGACGAGTGGCAGAGC 181

QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 182 GGCCAGCGCTGGGAACTGGCACTGGTCTGTTGGATTACCTCGCGTGGGAGACA 241

QY 61 LeuSerGluGlnValaGlnGluLeuLeuSerSerGlnValaThrGlnGluLeuArgAla 80
DB 242 CTGTCTGACAGCTGCGAGAGAGAGCTGCTCAGCTCCCAAGTCAACCAAGACTGAGGGCG 301

QY 81 LeuMetaspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
DB 302 CTGATGAGCAGAGACCATGAAAGAGTTGAAGCGCTCAAAATCGAAGTGGAGACAACTG 361

QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaIaGlnAla 120
DB 362 ACCCGGAGTGGAGAGAGCGGCGACGCTGCTCAAGAGACTCGACAGCGCGAGGCG 421

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValaGlnTyrArgGlyVala 140
DB 422 CGCGTGGGCGGACATGAGAGAGCGTGGCGGCGCTGCTGCTACCTACCGCGGCGAGGTG 481

QY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValaArgLeuAlaSerHisLeuArg 160
DB 482 CAGGCGCATGCTCGCGCAGAGACCGAGGAGCTGCGGCTGCGCTCCACCTGCGGC 541

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValaTyr 180
DB 542 AAGCTGCGTAAGCGGCTCTCCGCAATCCGATGACTGCAAGAGCGCTGCGACTGTAC 601

QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 602 CAGGCGGGGCGCGGAGGCGCGAGCGGCGCTCAGCGCATCCGCGAGGCGCTGCGG 661

QY 201 ProLeuValaGluGlnGlyArgValaArgAlaAlaThrValaGlySerLeuAlaGlyInPro 220
DB 662 CCCCTGTGTGAACAGGCGCGCTGCGGCGCGCTGCGCGCGGATGAGAGATGGGCG 721

QY 221 LeuGlnGluArgAlaGlnAlaThrPcLysGluArgLeuArgAlaArgMetGluGlnMetGly 240
DB 722 CTACAGAGAGCGGCGCGAGGCGCTGCGGCGCGCTGCGCGCGGATGAGAGATGGGCG 781

QY 241 SerArgTrpArgAspArgLeuAspGluValaLysGluGlnValaIaGlnValaArgAlaIys 260
DB 782 ACTCGAGACCGCGGACCGCTGCGAGAGGAGGAGGAGAGAGTGGCGGAGTGGCGCCAAAG 841

QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAla 277
DB 842 CTGGAGAGAGCGAGCCAGAGTACGCTGCGAGCGCGAGGCGCTTCCAGAGCGC 892

RESULT 10
AX409597
LOCUS AX409597
DEFINITION Sequence 2244 from Patent W00229103.

ACCESSION AX409597
VERSION AX409597.1 GI:21442302
KEYWORDS
SOURCE
ORGANISM human.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
AUTHORS Alvarres,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2244 11-Apr-2002;
GENE LOGIC INC (US)
FEATURES
source location/Qualifiers
1..1157
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="EMBL/Genbank Accession No. M12529"

BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN

Alignment Scores:
Pred. No.: 7 62e-72 Length: 1157
Score: 1365.00 Matches: 275
Percent Similarity: 99.28% Conservative: 0
Best Local Similarity: 99.28% Mismatches: 2
Query Match: 99.34% Indels: 0
DB: 6 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x AX409597 (1-1157)

QY 1 MettysValleuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20
DB 62 ATGAAGGTTCTGTGGGCTGCTGTGGTGCATCTCTGCGAGAGTCCAGGCCAAGGTG 121

QY 21 GluGlnAlaValaGluThrGluProGluLeuValGlnThrGluTrpGlnSer 40
DB 122 GAGCAAGCGGTGAGACAGAGCGGAGCCGAGCTGCGCAGACAGACGAGTGGCAGAGC 181

QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 182 GGCCAGCGCTGGGAACTGGCACTGGTCTGTTGGATTACCTCGCGTGGGAGACA 241

QY 61 LeuSerGluGlnValaGlnGluLeuLeuSerSerGlnValaThrGlnGluLeuArgAla 80
DB 242 CTGTCTGACAGCTGCGAGAGAGAGCTGCTCAGCTCCCAAGTCAACCAAGACTGAGGGCG 301

QY 81 LeuMetaspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
DB 302 CTGATGAGCAGAGACCATGAAAGAGTTGAAGCGCTCAAAATCGAAGTGGAGACAACTG 361

QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaIaGlnAla 120
DB 362 ACCCGGAGTGGAGAGAGCGGCGACGCGCTGCTCAAGAGACTGCAACAGCGCGAGGCG 421

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValaGlnTyrArgGlyVala 140
DB 422 CGCGTGGGCGGACATGAGAGAGCTGTGCGCGCTGCTGCTACCTACCGCGGCGAGGTG 481

QY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValaArgLeuAlaSerHisLeuArg 160
DB 482 CAGGCGCATGCTCGCGCAGAGACCGAGGAGCTGCGGCTGCGCTCCACCTGCGGC 541

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValaTyr 180
DB 542 AAGCTGCGTAAGCGGCTCTCCGCAATCCGATGACTGCAAGAGCGCTGCGACTGTAC 601

QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 602 CAGGCGGGGCGCGGAGGCGCGAGCGGCGCTCAGCGCATCCGCGAGAGCGCTGCGG 661

QY 201 ProLeuValaGluGlnGlyArgValaArgAlaAlaThrValaGlySerLeuAlaGlyInPro 220
DB 662 CCCCTGTGTGAACAGGCGCGCTGCGGCGCGCTGCGCGCGGATGAGAGATGGGCG 721

OY 221 LeuGlIngluArGlaInAlaInArGpLgLuArGleuArGAlaArGmetGluInMetGly 240
 DB 722 CTCACGAGAGGGGCCAGGCTGGGGGAGAGCGCGCGGATGAGAGATGGGC 781
 OY 241 SerArGThrArGAspArGleuAspGluValLysGluGlnValAlaGluValAlaGlyAlaLys 260
 DB 782 AGTCGAGACCGCGACCGCTGGAGAGGTGAAGAGCAGGTGGAGGAGGCGCGCCAG 841
 OY 261 LeuGlIngluGlnAlaGlnInAlaArGleuGlnAlaGlnAlaPheGlnAla 277
 DB 842 CTCGAGAGAGGAGCCACGACAGATACGCTTCAGGCGGAGGCTTCCAGGCC 892
 RESULT 11
 LOCUS 115975 1157 bp DNA linear PAT 03-APR-1996
 DEFINITION Sequence 1 from patent US 5472858.
 ACCESSION 115975
 VERSION 115975.1 GI:1250883
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1157)
 AUTHORS Attle,A.D., Gretch,D.G., Sturley,S.L. and Beckage,N.E.
 TITLE Production of recombinant proteins in insect larvae
 JOURNAL Patent: US 5472858-A 1 05-DEC-1995;
 FEATURES
 source 1..1157
 location/Qualifiers
 BASE COUNT 212 a 370 c 426 g 149 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 7.62e-72 Length: 1157
 Score: 1365.00 Matches: 275
 Percent Similarity: 99.28% Conservative: 0
 Best Local Similarity: 99.28% Mismatches: 2
 Query Match: 99.34% Indels: 0
 DB: Gaps: 0
 US-09-827-854-15_COPY_1_277 (1-277) x 115975 (1-1157)
 OY 1 MetLysValLeuTrpAlaAlaLeuLeuValTrpPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 62 ATGAAGTTCTGTGGCTGGCTGTGTGCATCTCTGCGACAGATGCCAGGCCAAGTG 121
 OY 21 GluGlnAlaValAlaGluTrpGluProGluProGluLeuArgGlnGlnTrpGluInser 40
 DB 122 GAGCAAGCGGTGAGACAGACGCCGAGCGGAGCTGCGCAGACCGAGTGGCAGAGC 181
 OY 41 GlyGlnArGTrpGluLeuAlaLeuGlyArGpHeTrpAspTrpLeuArGTrpValGlnTrp 60
 DB 182 GCGCAGCGCTGGGAAGTGGACACTGGGTCTTTGGGATTACTGCGCTGGCGTGCAGACA 241
 OY 61 LeuSerGluGlnValAlaGlnGluLeuLeuSerSerGlnValTrpGlnGluLeuArgAla 80
 DB 242 CTGTCTGACGAGGTGACGAGGAGCTGCTCACTCCCAAGTCAACCAAGAACTGAGGGCG 301
 OY 81 LeuMetAspGluTrpMetLysGluLeuLysAlaTrpLysSerGluLeuGluGluInleu 100
 DB 302 CTCATGACGACGACCAATGAAGGAGTTGAAGGCTTACAAATCGAATGGAGTGAAGCAACTG 361
 OY 101 ThrProValAlaGluGluTrpArGAlaArGLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 362 ACCCGCGTACGAGAGAGACGCGGCGACGCTCCAGGAGCTGCAGACGGCGCAGGCC 421
 OY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
 DB 422 CGGCTGGGGCGGACATGAGGAGCGTGTGGCGCGCTGTGTGACGAGTACCGCGCGAGGTG 481
 OY 141 GlnAlaMetLeuGluInserThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 161 LysLeuArGLeuArGleuLeuArGAspAlaAspAspLeuGlnLysArGleuAlaValTrp 180
 DB 542 AAGCTCGTAAGAGGCGTCTCCGCGATCCCGATGACCTCGACAGAGCGCTGGCAGTGC 601
 OY 181 GlnAlaGlyAlaArGluGluGluValArgGlyLeuSerAlaAlaArgGluArgLeuGly 200
 DB 602 CAGCGCGGGGCCCGCAGGCGCGCGCGCGCTTCAGGCGCCATCCGACGAGCGCTGGGG 661
 OY 201 ProLeuValGluGlnGluArgValAlaArgAlaAlaTrpValGlySerLeuAlaGluInleu 220
 DB 662 CCCCTGTGAGACAGAGGCGCGCTGCGCGCGCGCTTCAGGCGCTTCAGGCGCGCGCG 721
 OY 221 LeuGlIngluArGlaInAlaInArGpLgLuArGleuArGAlaArGmetGluInMetGly 240
 DB 722 CTCACGAGAGGGGCCAGGCTGGGGGAGAGCGCGCGGATGAGAGATGGGC 781
 OY 241 SerArGThrArGAspArGleuAspGluValLysGluGlnValAlaGluValAlaGlyAlaLys 260
 DB 782 AGTCGAGACCGCGACCGCTGGAGAGGTGAAGAGCAGGTGGAGGAGGTGGCGCAG 841
 OY 261 LeuGlIngluGlnAlaGlnInAlaArGleuGlnAlaGlnAlaPheGlnAla 277
 DB 842 CTCGAGAGAGGAGCCACGACAGATACGCTTCAGGCGGAGGCTTCCAGGCC 892
 RESULT 12
 HUMAPOE 1157 bp mRNA linear PRI 08-AUG-1995
 LOCUS Human apolipoprotein E mRNA, complete cds.
 DEFINITION M12529
 ACCESSION M12529.1 GI:178848
 VERSION M12529.1 GI:178848
 KEYWORDS
 SOURCE Homo sapiens (clone: PHAE[112,178,813].) male 57-year old liver
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1157)
 AUTHORS McLean,J.W., Elshoutday,N.A., Chang,D.J., Mahley,R.W. and
 Taylor,J.M.
 TITLE Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing
 of a new variant
 JOURNAL J. Biol. Chem. 259 (10), 6498-6504 (1984)
 MEDLINE 84212473
 PUBMED 6327682
 FEATURES
 source 1..1157
 location/Qualifiers
 1..1157
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="19q13.2"
 /clone="PHAE[112,178,813]."
 /sex="male"
 /tissue_type="liver"
 /dev_stage="57-year old"
 1..1157
 /gene="APOE"
 <1..1157
 /note="APOE"
 /note="APOE mRNA; 600-119-691"
 62..1015
 /gene="APOE"
 /note="precursor"
 /codon_start=1
 /product="apolipoprotein E"
 /protein_id="AAB59518.1"
 /db_xref="GI:178849"
 /db_xref="GDB:600-119-691"
 /translation="MKVLMAALLVFLFLGCOAKVEQAVETPEPELROOTMOSGORM
 ELAIGRPMDYLRMVOITLSEVOEELLSQVQELALNDETMKELKAVKSELEQITP
 VAEETRARLSKEIOTFAORLIGADMGVGRIVYQGEVQAMIGOSTEELRVLRLSHLR
 KLRRLDPPDDLRRLAVYQAGAREGEFRGISATREILGVLVEQGRVRAATVYSILAG
 OPLERRAOWBERLRRMEWGSRTDRLEDEVKQVAVRAKLEBOAOQIRLOAEAFQ

sig_peptide ARKSWFEPLVEDMOROMAGLVEKVAAGTSAAPVPSDNH^{62..115}
 /gene="APOE"
 /note="G00-119-691"

mat_peptide 116..1012
 /gene="APOE"
 /product="apolipoprotein E"
 /note="G00-119-691"

variation 370
 /gene="APOE"
 /note="a in pHAEl112,178]; g in pHAEB13; G00-119-691"

variation 410
 /replace="g"
 /gene="APOE"
 /note="a in pHAEl112,178]; g in pHAEB13; G00-119-691"

variation 569
 /replace="g"
 /gene="APOE"
 /note="c in pHAEl112,178]; g in pHAEB13; G00-119-691"

variation 784
 /replace="g"
 /gene="APOE"
 /note="c in pHAEl112,178]; g in pHAEB13; G00-119-691"

variation 859
 /replace="c"
 /gene="APOE"
 /note="g in pHAEl112,178]; a in pHAEB13; G00-119-691"

BASE COUNT 212 a 370 c 426 g 149 t
 ORIGIN 30 bp upstream of BstNI site.

Alignment Scores:
 Pred. No.: 7.62e-72 Length: 1157
 Score: 1365.00 Matches: 275
 Percent Similarity: 99.28% Conservative: 0
 Best Local Similarity: 99.28% Mismatches: 2
 Query Match: 99.34% Indels: 0
 DB: 9 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x HUMAPOE (1-1157)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||
 DB 62 ATGAGGTTCTGTGGCGTGGCTGCTGTCGACATTCCTGCGAGATGCCAGATGCCAAGGTC 121

QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 |||
 DB 122 GAGCAAGCGGTGAGACACAGCCGAGCCGAGCTGCGCAGCAGACGACGAGCGAGAGC 181

QY 41 GlyLnaArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 |||
 DB 182 GGCACAGCGCTGGAACTGGCACTGGCTCTTTGGGATTACCTGCGCTGGGTGCAGACA 241

QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||
 DB 242 CTGTGTGACAGAGTGCAGAGAGCTGCTCAGCTCCCAAGTCAACCAAGAACTGAGAGGCG 301

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu 100
 |||
 DB 302 CTGATGAGACAGACCATGAAAGAGTTGAAGGCTCAATCGAACTGGAGAGAAACATCG 361

QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaLysVal 120
 |||
 DB 362 ACCCGGTGAGCGAGAGAGCGGCGACGCTGTCCAAAGAGACTGCAGAGCGCGACAGCC 421

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 |||
 DB 422 CGGCTGGGCGCGAGACATGAGAGAGCTGTGCGCGCTGCTGTCAGTACCGCGGAGGTG 481

QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||
 DB 482 CAGGCATCTCGGCGAGAGACCGAGAGAGCTGCGGCTGCTGCTCCACCTCCGCGC 541

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||
 DB 542 AAGCTGGCGTAACGGGCTCTCCGGCATCCCGATGACCTGACAAACAGCGCTGGCACTGTAC 601

QY 181 GlnAlaGlyAlaArgGluGluValaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 |||
 DB 602 CAGGCGGGGGCCCGGAGAGGCGCCAGCGCGGCTCAGCGCCATCCGAGAGCGCTGGGG 661

QY 201 ProLeuAlaGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 |||
 DB 662 CCCCTGGTGGACAGGCGCGCTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721

QY 221 LeuGlnGluArgAlaGlnAlaThrPglGluArgLeuArgAlaArgMetGluGluMetGly 240
 |||
 DB 722 CTACAGAGCGGGCCAGGCGCTGGGGCCAGCGGCTGGCGCGCGATGAGAGATGGCG 781

QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValaGluValaArgAlaLys 260
 |||
 DB 782 ACTCGAGCCCGGACCGCGCTGAGCAGAGTGAGAGAGAGAGTGGCGAGGTGGCGCCCAAG 841

QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAla 277
 |||
 DB 842 CTGAGAGAGACAGGCCAGCAGATACGCTGACAGCGCGAGCCCTTCCAGGCC 892

RESULT 13
 LOCUS BD004279 1156 bp DNA linear PAT 31-JAN-2002
 DEFINITION Apo E humanized mammal.
 ACCESSION BD004279.1 GI:18632240
 VERSION BD004279.1
 KEYWORDS JP 2001017028-A/3.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1156)
 Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.
 Apo E humanized mammal
 Patent: JP 2001017028-A 3 23-JAN-2001;
 JOURNAL MITSUBISHI CHEMICAL CORP
 OS Homo sapiens (human)
 PN JP 2001017028-A/3
 PD 23-JAN-2001
 PF 28-APR-2000 JP 2000128919
 PR

COMMENT
 PI SHINOBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA PC
 A0167/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10, PC
 C12N15/09//C07K14/775,
 PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,(C12N5/00,C12R1:91) CC

FEATURES
 source location/Qualifiers
 FT Key CDS location/Qualifiers
 FT CDS (61)..(1011).

BASE COUNT 208 a 369 c 432 g 147 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.14e-71 Length: 1156
 Score: 1362.00 Matches: 276
 Percent Similarity: 99.64% Conservative: 0
 Best Local Similarity: 99.64% Mismatches: 1
 Query Match: 99.13% Indels: 0
 DB: 6 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x BD004279 (1-1156)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||
 DB 61 ATGAGGTTCTGTGGCGTGGCTGCTGTCGACATTCCTGCGAGATGCCAGATGCCAAGGTC 120

QY	21	GIUAGNAlaValaIGlunhrGluuprGluuPprGluuAArgInglunhrGluuPpInser	40
QY	121	GGAGCAAGCGGTGGAGACAGAGCCGGAGCCCGAGCTCCGACAGACCCAGTGGCGAGAC	180
QY	41	GIYGIunArGTrpGluLeuAlaLeuGIuArGpHeTrpAspTyrLeuArGTrpValGIunhr	60
Db	181	GGCCACCGCTGGGAACCTGGACACTGGAGGTGGCTTTTGGAGATTACCTGGCGTGGGGACAGACA	240
QY	61	LeuSerGIuGIunValGIunGIuLeuLeuSerSerGIunValThrGIunGIuLeuArGAla	80
Db	241	CGTCTGACAGAGGTGAGAGAGAGCGTCACGTCTCCAGCTCACCCAGAACACGAGGGGG	300
QY	81	LeuMetAspGIunThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIunGIunGIunLeu	100
Db	301	CGATGACAGACCACTGAAGAGTTGAAGGCTTACAAATTCGAACTGAGAGGAACACATG	360
QY	101	ThrProValAlaGIuGIunThrArGAlaArGLeuSerLysGIuLeuGIunAlaAlaGIunAla	120
Db	361	ACCCGGTGGCGAGAGAGACCGCGGACCGGCTGTCCAGAGAGCTGCAAGCGCGGCGAGGCC	420
QY	121	ArgLeuGIuAlaAspMetGIunAspValCysGIuArgLeuValGIunTyrArgGIuGIunVal	140
Db	421	CGGCTGGCGCGACATGAGAGAGCGTGGCGCGCGCTGGTGGACATGACCGGGCGAGGTG	480
QY	141	GIunAlaMetLeuGIunGIunSerThrGIuGIuLeuArGValArgLeuAlaSerHisLeuArg	160
Db	481	CAGGCCATCTCTGGCCAGACACCGAGAGATCGGGGTCCGCTCCCTCCACACGTGGCC	540
QY	161	LysLeuArgLysArGLeuLeuArGAspAlaAspAspLeuGIunLysArgLeuAlaValTyr	180
Db	541	AAGCTGCTGACAGCGCTCTCCGCGATGCCATGACCTCCAGAACGCGCTGGCAGGTGAC	600
QY	181	GIunAlaGIuAlaArGLeuGIunGIuAlaGIuArGArgLysSerAlaIleArGLeuArgLeuGIu	200
Db	601	CAGCGCGGGGCCCGAGAGGGCGCGGAGCGCGGGCTCAGCGCCATCCGAGCGGCTGGGG	660
QY	201	ProLeuValGIuGIunGIuArgLysArgValArgAlaAlaThrValGIuSerLeuAlaGIuInPro	220
Db	661	CCCTGTGTGAACAGGCGCGGTGGCGGGCGGCACACTGTGGCTCCCTGGCGGGCCAGCCG	720
QY	221	LeuGIunGIuArGAlaGIunAlaTrpGIuGIuArgLeuArGAlaArgMetGIuGIuMetGIu	240
Db	721	CTACAGAGAGCGGGCCAGGCTCTGGGGCGAGCGGCTGTGGCGCGGATGGAGAGATGGGG	780
QY	241	SerArgThrArgAspArgLeuAspGIuValLysGIuGIunValAlaGIuValArgAlaLys	260
Db	781	ACCCGGAGCCCGGAGACCGCTCGACAGCGGTGAAGAGAGACAGTGGCGAGGTGGCGCGCAG	840
QY	261	LeuGIuGIuGIunAlaGIunGIuLeuArgLeuGIunAlaGIuAlaPheGIunAla	277
Db	841	CTGGAGGAGACAGGCCAGACGATAGCCCTGGCAGGCCGAGGCTTCACAGGCC	891
RESULT 14			
MAPOE			
LOCUS	MAPOE	1178 bp	mRNA
DEFINITION	Monkey mRNA for apolipoprotein E.		linear
ACCESSION	X13887		
VERSION	X13887.1	GI:38054	
KEYWORDS	apolipoprotein; apolipoprotein E.		
SOURCE	Macaca fascicularis.		
ORGANISM	Macaca fascicularis		
REFERENCE	Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteleostomi; Primates; Catarrhini; Cercopithecoidea; Cercopithecidae; 1 (bases 1 to 1178)		
AUTHORS	Marotti,K.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-DEC-1988) Marotti K.R., The Upjohn Co, 7242 209 713,		
REFERENCE	301 Henrietta Street, Kalamazoo, MI 49008		
AUTHORS	Marotti,K.R., Whitted,B.E., Castle,C.K., Polites,H.G. and Melchior,G.W.		
TITLE	Nucleotide sequence of the cynomolgus monkey apolipoprotein E cDNA		

Journal	Nucleic Acids Res.	17 (4)	1778 (1989)
Medline	89160349		
PubMed	2922300		
Comment	Data kindly reviewed (20-Mar-1989) by Marotti K.R.		
Features	Location/Qualifiers		
Source	1..1178		
	/organism="Macaca fascicularis"		
	/db_xref="taxon:9541"		
	/clone="1"		
	/rname="type="11ver"		
	/clone_lib="lambda gt11"		
	83..1036		
CDS	/codon_start=1		
	/product="proapolipoprotein E (AA -18 to 299)"		
	/protein_id="CAA32092.1"		
	/db_xref="GI:38055"		
	/db_xref="SWISS-PROT:P10517"		
	/translation="MKVLMALVTPFALGKQAKVEQVPEPTEPDLRQAEQSGQPM ELAGRFMDYRWYQTSLEVOQELLSPOYQELTMTDMETKAKSELEQSLP VAEETRARLSEKLEDAQARIGADMEDYRSRLVQTRSVQMLQSTELRLASHLR KLRLRLDADLDQRLAVYQAGAREAEAGVSAIRRLQPLVEQGRVATVGSLAS QPLDERAQLGEERLRARMEKSGRTLRRLDEVOAEVRAKLEEQAOQTSLEAEAFQ ARLSWEPLEVEDMQRQWAGLVEKVEQVAAVASTAPVVDNH"		
	83..136		
sig_peptide	137..1033		
mat_peptide	/product="apolipoprotein E (AA 1 - 299)"		
misc_feature	1154..1159		
	/note="polyA signal"		
BASE COUNT	217 a 373 c 426 g 162 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	2,12e-65	Length:	1178
Score:	1255.00	Matches:	259
Percent Similarity:	94.22%	Conservative:	2
Best Local Similarity:	93.50%	Mismatches:	16
Query Match:	91.34%	Indels:	0
DB:	9	Gaps:	0
US-09-827-854-15_COPY_1_277 (1-277) x MFAPOE (1-1178)			
QY	1 MetLysValLeuThrPAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20	
Db	83 ATGAAGGTTCTGTGGGCGCTGCTGCTGCATTCCTGGAGAGATGCCAGGCCAAGTG	142	
QY	21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer	40	
Db	143 GAGCAACCGGTGGAGCCACAGACAGCAACCCAGCTTCCCGACGACGCGAGGCCAGAGC	202	
QY	41 GlyIleArgTrpGluLeuAlaLeuGluYlArgPheTrpAspTrpLeuArgTrpValGlnThr	60	
Db	203 GGGCAGCCCTGGAGACTGGCACTGGGTGGCTTTGGGATTACCTGGCGCTGGTGACAGCA	262	
QY	61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80	
Db	263 CTGTCTGAGCAGGTGCACGAGAGAGGTGCTCAGCCCGCAGGTCCACCCAGGAATGACGACG	322	
QY	81 LeuMetAspGluThrMetLysGluLeuMetLysAlaTrpLysSerGluLeuGluGlnGluLeu	100	
Db	323 CTGATGGACGAGACCATAGAGAGGTGAAGGCTTACCAATGGAACATGGAGAACAGTG	382	
QY	101 ThrProValAlaGluGluThrArgAlaIleArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120	
Db	383 AGCCCGGGGGGAGAGAACGCGGGGACGCGCTGTGCCAAGAGAGCTGACAGCGCGCCAGGCC	442	
QY	121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyValAl	140	
Db	443 CGGCTGGGTGCGCAGCATGGAGAGACTGGCGCAGCCGCTGGGTGCAAGTACCAGCAGAGGTG	502	
QY	141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaIleSerHisLeuArg	160	
Db	503 CAGGCATGTGGGGCCAGATGACGAGAGAGGTGGGGGGCGGCTGTGGCTCCACACTGGCC	562	

QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	563	AAGCTGGCAGAGGGCTCTCCGGATGCTGATGACCTGCAGAGAGCGCTGGCAGTGTAT	622
QY	181	GlnAlaGlyAlaArgGlyAlaGlyAlaGlyLeuSerAlaIleArgGlyArgLeuGly	200
Db	623	CAGCCCGGGCCCGCAGGGCGCGGAGCCGGGGTCAAGCCCATCCGCGAGCGCTGGGA	682
QY	201	ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro	220
Db	683	CCCTGGGTGAGCAGGGCCCGTGGGGCCGCCACCTGTGGCTCCCTCCGACCCAGCCG	742
QY	221	LeuGlnGluArgAlaGlnAlaTrpGlyGlyArgLeuArgAlaArgMetGlnGluMetGly	240
Db	743	CTTCAGAGAGGGGCCAGGCTTGGGTGAGCGGGCTTCGCCACAGATGGAGATGGGC	802
QY	241	SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlyAlaArgAlaLys	260
Db	803	AGCCGGACCGCGACCGCTGGACGAGGTAGAGAGAGGTGGCGGGGTGCCGCCAAG	862
QY	261	LeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	277
Db	863	CTGAGAGAACAGGCCACGACATAGCTGCAGGCCGAGGCTTCCAGGCC	913
RESULT	15		
LOCUS	AF261279	5491 bp	DNA linear PRI 27-OCT-2000
DEFINITION		Homo sapiens apolipoprotein-E gene, complete cds.	
ACCESSION	AF261279		
VERSION	AF261279.1	GI:11034800	
KEYWORDS			
SOURCE		Homo sapiens.	
ORGANISM		Homo sapiens	
REFERENCE			
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
TITLE		Nickerson,D.A., Taylor,S.L., Fullerton,S.M., Weiss,K.M.,	
JOURNAL		Clark,A.G., Stengard,J.H., Salomaa,V., Boerwinkle,E. and Sing,C.F.	
		Sequence diversity and large-scale typing of SNPs in the human	
		apolipoprotein E gene	
JOURNAL		Genome Res. 10 (10), 1532-1545 (2000)	
MEDLINE		20499366	
PUBMED		11042151	
REFERENCE		2 (bases 1 to 5491)	
AUTHORS		Nickerson,D.A.	
TITLE		Direct Submission	
JOURNAL		Submitted (27-APR-2000) Department of Molecular Biotechnology,	
		University of Washington, Box 357730, Seattle, WA 98195, USA	
FEATURES			
source		1. 5491	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/chromosome="19"	
		/map="19q13.2"	
		<3. .>108	
repeat_region		/note="putative"	
		/rpt_family="MIR"	
		/rpt_type-dispersed	
		73	
variation		/frequency="0.01"	
		/replace="t"	
repeat_region		<207. .>295	
		/note="putative"	
		/rpt_family="MIR"	
		/rpt_type-dispersed	
		308	
variation		/frequency="0.01"	
		/replace="t"	
		<320. .>339	
satellite		/note="putative"	
		/rpt_type-tandem	
		<340. .>637	
repeat_region		/note="putative"	

		/rpt_family="Alu"	
		/rpt_type-dispersed	
		471	
variation		/frequency="0.01"	
		/replace="g"	
		545	
variation		/frequency="0.01"	
		/replace="t"	
		560	
variation		/frequency="0.22"	
		/replace="t"	
		624	
variation		/frequency="0.07"	
		/replace="c"	
		<638. .>718	
satellite		/note="putative"	
		/rpt_type-tandem	
		832	
variation		/frequency="0.45"	
		/replace="t"	
		join(1060, 1094, 1855, 1920, 3013, 3205, 3786, 4645)	
mRNA		/product="apolipoprotein-E"	
variation		1163	
		/frequency="0.35"	
		/replace="c"	
		1522	
variation		/frequency="0.01"	
		/replace="a"	
		1575	
variation		/frequency="0.01"	
		/replace="t"	
		join(1878, 1920, 3013, 3205, 3786, 4503)	
CDS		/note="APOE"	
		/codon_start=1	
		/product="apolipoprotein-E"	
		/protein_id="AAG27089.1"	
		/db_xref="GI:11034801"	
		/translation="MKVLAALVFLAGCOAKKVAQAVETPEPELROTEMSGORM	
		ELALGRFMDYLRWVQTLSEQNOEILSQVDELRALNDETKEKAKKSELDLTP	
		VAETPRARKSLQAQAARLQADMDVCGRLVQNGEYQAMLGSTELRLVLSHLR	
		KLRRLRLDADDLQRLAVYQAGAREGELSAIRERGLPVEGRVRAVGLAG	
		QPLQRAQAWGERLRARMEQMSRTDRIDEVKEQVAEVRARLEQAOQIRLQAEAFQ	
		ARLKSFEFLVEDMORWAGLVEKVOAAVGTSAAPVPSDNI"	
		1998	
variation		/frequency="0.10"	
		/replace="a"	
		<2124. .>2435	
repeat_region		/note="putative"	
		/rpt_family="Alu"	
		/rpt_type-dispersed	
		2440	
variation		/frequency="0.21"	
		/replace="a"	
		<2569. .>2848	
repeat_region		/note="putative"	
		/rpt_family="Alu"	
		/rpt_type-dispersed	
		2907	
variation		/frequency="0.01"	
		/replace="g"	
		3106	
variation		/frequency="0.01"	
		/replace="c"	
		<3472. .>3588	
repeat_region		/note="putative"	
		/rpt_family="Alu"	
		/rpt_type-dispersed	
		3673	
variation		/frequency="0.01"	
		/replace="g"	
		3701-3702	
variation		/frequency="0.01"	
		/replace="ct"	

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:08:17 ; Search time 156.176 Seconds
(without alignments)
3994.237 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_277
1374
Sequence: 1 MKVLMALVLTFLAGCOAKV.....RAKLEQAQQRIGQAEAFQA 277

Scoring table:
BLSDM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+._p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09827854/funat_11032003_101609_27476/app_query.fasta_1.3576
-DB=N.Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blscorem2 -TRANS=human40.cdl
-LIST=45 -DOCCALIGN=200 -THR.SCORE=oct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854.ecgn1.1-1201.etrnat_11032003_101609_27476 -NCPU=6 -ICPU=3
-NO_XLPPY -NO_MMAP -LARGEQUERY -NES-SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.Geneseq_101002:*

1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1374	100.0	954	24 AAD26035	Human apolipoprotein E
2	1374	100.0	1110	7 AAN60409	Human apolipoprotein E
3	1374	100.0	1147	22 ABA83113	Human apolipoprotein E
4	1374	100.0	1156	22 AAF84315	Human apolipoprotein E
5	1374	100.0	1156	24 AAD22048	Human apolipoprotein E
6	1370	99.7	1156	24 AAD22052	Human apolipoprotein E
7	1366	99.4	1156	22 AAF84314	Human apolipoprotein E
8	1366	99.4	1156	24 AAD22049	Human apolipoprotein E
9	1366	99.4	1156	24 AAD22051	Human apolipoprotein E
10	1365	99.3	1157	17 AAT06957	Human apolipoprotein E
11	1365	99.3	1157	24 ABA84514	Human apolipoprotein E
12	1365	99.3	1157	24 ABA84514	Human apolipoprotein E
13	1365	99.3	1157	24 ABA84514	Human apolipoprotein E
14	1364	99.3	1110	6 AAN50450	Sequence encoding
15	1362	99.1	1156	22 AAF84316	Human ApoE coding
16	1362	99.1	1156	24 AAD22047	Human apolipoprotein E
17	1359	98.9	1156	24 AAD22050	Human apolipoprotein E
18	1322	96.2	1279	22 AAS22437	Human cDNA encoding
19	1251.5	91.1	1107	19 AAX75756	Human apolipoprotein E
20	1185.5	86.3	9360	24 ABL31915	Human ApoE genomic
21	1185.5	86.3	10716	24 AAD26034	Human apolipoprotein E
22	1180	85.9	3805	20 AAZ09524	Human Apo E genomic
23	1180	85.9	3805	20 AAZ09526	Human Apo E genomic
24	1158.5	84.3	10716	24 AAD26108	Human Apo E genomic
25	1004.5	73.1	965	24 AAD32081	Bovine ApoE gene.
26	1004.5	73.1	5617	24 AAD32077	Human alpha-1-anti
27	1004.5	73.1	6026	24 AAD32075	Human albumin prom
28	1002.5	73.0	1126	19 AAV29159	Nucleotide sequence
29	984	71.6	660	18 AAT69792	Partial human apol
30	919	66.9	936	17 AAO69101	ApoE1L2 protease
31	919	66.9	936	17 AAT18070	Human ApoE1L2 cDN
32	881	64.1	597	17 AAT18068	Human ApoE4 cDNA.
33	872	63.5	597	15 AAO69099	ApoE1L protease cD
34	823.5	59.9	1381	22 AAS22673	Human cDNA encoding
35	678	49.3	600	20 AAV89595	EST clone Cpl47.
36	656.5	47.8	786	21 AAF18114	Lung cancer associ
37	651	47.4	407	24 ABR34238	Human cDNA for nov
38	625	45.5	478	24 ABR95994	Gene #2492 used to
39	625	45.5	478	24 ABL2679	Colon adenocarcino
40	625	45.5	478	24 ABL67340	Thyroid cancer rel
41	594	43.2	499	22 AAS03049	Human diagnostic a
42	514	37.4	330	12 AAO11980	Human apolipoprote
43	504	36.7	405	21 AAC02139	Human secreted pro
44	485	35.3	345	22 AAB98479	Human EST-derived
45	419	30.5	260	21 AAA40342	Human apolipoprote

ALIGNMENTS

RESULT 1
AAD26035
ID AAD26035 standard; cDNA: 954 BP.

AAD26035:
26-MAR-2002 (first entry)

Human apolipoprotein E (APOE) cDNA.

Human: antiLipemic; neuroprotective; nootropic; genetic variant; APOE;
apolipoprotein E; haplotyping; familial dysbetalipoproteinemia; therapy;
genotyping; type III hyperlipoproteinemia; Alzheimer's disease; SNP;
atherosclerosis; single nucleotide polymorphism; chromosome 19q13.2; ss.

Homo sapiens.
Location/Qualifiers
key

PT	CDS	1..954
FT	/**tag= a	
FT	/product= "Human APOE protein"	
FT	replace (13, C)	
FT	/**tag= a	
FT	/standard_name=	"Single nucleotide polymorphism (SNP) "
FT	replace (31, G)	
FT	/**tag= b	
FT	/standard_name=	"Single nucleotide polymorphism (SNP) "
FT	replace (364, A)	
FT	/**tag= c	
FT	/standard_name=	"Single nucleotide polymorphism (SNP) "
FT	replace (388, C)	
FT	/**tag= d	
FT	/standard_name=	"Single nucleotide polymorphism (SNP) "
FT	replace (498, G)	
FT	/**tag= e	
FT	/standard_name=	"Single nucleotide polymorphism (SNP) "
FT	replace (526, T)	
FT	/**tag= f	
FT	/standard_name=	"Single nucleotide polymorphism (SNP) "
FT	replace (622, A)	
FT	/**tag= g	
FT	/standard_name=	"Single nucleotide polymorphism (SNP) "
XX	MO200179234-A2.	
PD	25-OCt-2001.	
XX		
XX	16-Apr-2001; 2001WO-US12303.	
PF		
XX	14-Apr-2000; 2000US-197188P.	
PR		
XX		
PA	(GENA-) GENAISSANCE PHARM INC.	
PL	Choi JY, Kilem SE, Koshy B, Lee HH;	
DR	WPI: 2002-075064/10.	
DR	P-PSDB; AAEL5158.	
PT	Genotyping human apolipoprotein gene of individual for determining haplotype of individual, involves determining identity of nucleotide pair at specific polymorphic sites for two copies of gene -	
PS	Claim 26; Fig 2; 78pp; English.	
CC	The patent discloses novel genetic variants of human apolipoprotein E (APOE) gene. The invention also relates to compositions and methods for haplotyping and/or genotyping the APOE gene. The haplotyping methods of the invention are useful for improving the efficacy and reliability of several steps in the discovery and development of drugs for treating diseases associated with APOE activity, e.g. familial dysbetalipoproteinaemia, type III hyperlipoproteinaemia, atherosclerosis, and Alzheimer's disease. They are useful to validate APOE as a candidate agent for treating a specific condition or disease predicted to be associated with APOE activity and in the design of clinical trials of candidate drugs for treating a specific condition or disease predicted to be associated with APOE activity. Genotyping or haplotyping methods are useful to screen for compounds targeting APOE to treat a specific condition or disease associated with APOE activity. The present sequence is a cDNA encoding human APOE protein. APOE gene is located on chromosome 19q13.2.	
SQ	Sequence 954 BP; 168 A; 290 C; 377 G; 119 T; 0 other:	
Alignment Scores:		
Pred. No.:	1,84e-102	Length: 954
Score:	1374..00	Matches: 277
Percent Similarity:	100.00%	Conservative: 0
Best local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
Gaps:	24	Gaps: 0
OB:		

US-09-827-854 -15_COPY_1_277 (1-277) x AND26035 (1-954)	
OY 1 MetlysValLeuTrpPalaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal	20
Db 1 ATGAAGGTTCTGTGGGCTGCTTCTGGTCAATTCTCTGGAGGATCCAGGCCAAGGTG	60
OY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuLeuArgGlnGlnThrGlnThrGlnThr	40
Db 61 GAGCAAGCGGTGGAGACAGAGACCCGAGCTGGCGCCAGCAGAGACGAGTGGCAGAGC	120
OY 41 GlyGlnArgTrpGlnLeuAlaAlaLeuGlyAraPheThrPaspTyrLeuArgTrpValGlnThr	60
Db 121 GGGCAGCGCTGGAAACTGGCACTGGTGCTTTGGGATTACGTGGCTGGTGCGAGACA	180
OY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla	80
Db 181 CTGTCTGTAGCAGGTGCAGGAGAGACTGCTCTAGCTCCAGGTCAACCCAGGAATGAGGGCG	240
OY 81 LeuMetAspGlnThrMetLeysGlnLeuLeuValaIaTyrIysSerGlnLeuGlnGlnGln	100
Db 241 CTGATGGACGAGACCATGAAGATTGAAGGCTTCAAAATCGGAAGCTGGAGGAACAATG	300
OY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerIysGlnLeuGlnAlaAlaGlnAla	120
Db 301 ACCCGCGTGGGGGAGAGAGACCGGGGCAAGGCTGTCCAAGAGAGCTGCAGCGGGCGAGGCC	360
OY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyAraGlnLeuValGlnTyrArgIylGlnVal	140
Db 361 CGGCTGGGCGCGGACATGGAGAGAGTGTGGCGGCTGTGGTGCAGTACCAGCGGAGAGT	420
OY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg	160
Db 421 CAGGCCATGCTGGCCAGAGACCCAGAGAGCTGGGGGTGGCTCGCTGCCACTGGCCG	480
OY 161 LysLeuArgIysArgLeuLeuArgAspAlaAspAspLeuGlnIlysArgLeuAlaValTyr	180
Db 481 AAGCTGCGCTAAGCGGCTCTCTCGGATGCCGATACCTGCGAAGAGCCCTCGCAGAGTAC	540
OY 181 GlnAlaGlnAlaArgGlnGlyAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly	200
Db 541 CAGGCGGGGCGCCGAGAGGCGCCGAGCGGCTCAGCCGCAATCCGCGCAGCGCCCTGGGG	600
OY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValAlaIysSerLeuAlaGlyGlnPro	220
Db 601 CCCCTGGGGAACAGGGCGCGGTGGGGCGCCACTGTGGGCTCCCTGGCGGGCCAGCGG	660
OY 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly	240
Db 661 CTACAGAGAGCGGGCCAGAGCCTGGGGGAGAGCGCTGCGCCCGCGGATGAGAGAGATGGGC	720
OY 241 SerArgThrArgAspArgLeuAspGlnValIysGlnGlnValAlaGlnValArgAlaIys	260
Db 721 AGCGGAGACCCCGACCCCTGTGAGAGAGTAAAGAGCAGGTGGCGGAGAGTGGCGCCCAAG	780
OY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleGlnAla 277	
Db 781 CTGAGAGAGCAGGCGCCAGCAGATACGCTCAGGCGCAGGCGCTTCCAGAGCC 831	
RESULT 2	
AAN60409 standard; DNA: 1110 BP.	
AAN60409;	
AAN60409;	
01-JAN-1980 (first entry)	
Human apolipoprotein-E.	
Apolipoprotein-E; hyperlipidemia; arteriosclerosis; ss.	
Homo sapiens.	
Key Location/Qualifiers	


```
FT CDS 15..968
FT /*tag= a
XX AU8547513-A.
XX 24-APR-1986.
XX 17-SEP-1985; 85AU-0047513.
XX 11-JUN-1985; 85JP-0126989.
XX (MITU ) MITSUBISHI CHEM IND KK.
XX WPI: 1986-150217/24.
XX P-PSDB: AAP60507.
XX
XX New DNA sequence coding for human apolipoprotein-E - and
XX expression vectors and transformed cells contg. It
XX
XX Disclousure: Fig 2; 45pp; English.
XX
XX The encoded protein is used to treat subjects who are deficient in
XX apolipoprotein-E (or who produce abnormal forms of this molecule)
XX and therefore are likely to suffer from hyperlipidemia, resulting in
XX arteriosclerosis. It can also be used to raise antisera for
XX detecting the protein deficiency or production of abnormal forms.
XX
XX Sequence 1110 BP; 198 A; 354 C; 415 G; 143 T; 0 other:
XX
XX Alignment Scores:
XX Pred. No.: 2.19e-102 Length: 1110
XX Score: 1374.00 Matches: 277
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX Db: 7 Gaps: 0
XX
XX US-09-827-854-15_COPY_1_277 (1-277) x AAN60409 (1-1110)
XX
XX 1 MettysValleuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
XX 15 ATGAAGGTTCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 74
XX
XX 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGln 40
XX 75 GAGCAAGGCGGTGGAGACAGACCGGAGCCGAGCTGCCCGACAGACCGAGTGG 134
XX
XX 41 GlyGlnAlaArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
XX 135 GGCCAGCGCTGGAACTGGCACTGGCTGCTTTGGGATTCACCTGCGCTGGGTG 194
XX
XX 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
XX 195 CTGCTGAGCAGGTGCAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 254
XX
XX 81 LeuMetAspGluThrMetLeuGluLeuLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
XX 255 CTGATGAGCAGACCATTAAGAGCTGAAGCCCTACAAATGGAATCGAAGCAACTG 314
XX
XX 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
XX 315 ACCCGGCGGCGAGAGACCGCGGCGCTGCTCAAGAGAGTCAGAGCGCGCAGGCC 374
XX
XX 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
XX 375 CGGCGGGGCGCCAGATGAGAGAGCTGTGCGCGCGCTGTGAGTACCGGCGAGGTG 434
XX
XX 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValAlaGlnAlaSerHisLeuArg 160
XX 435 CAGGCGATGCTCGGCGCAGAGCAGCAGAGAGCTGGGGTGGCCCTCCCTCCACTGGGC 494
XX
XX 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
XX 1110
```

```
Db 495 AAGCTGCGTAAGCGGCTCTCCCGCATGCCGATGACCTGCAGAACCGCTGGCACTGTAC 554
XX
XX 181 GlnAlaGlyAlaArgGluGluGluValArgGlyLeuSerAlaIleArgGluArgGly 200
XX 555 CAGGCGGGGCGCGCGAGGCGCGGAGCGGCTCAGCGCCATTCGCGAGCGCTGGGG 614
XX
XX 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
XX 615 CCCCTGTGTGAACAGAGCGCGGCTGGCGCGCCACTGTGTGGGTCTCTGGCGCGCAGCG 674
XX
XX 221 LeuGlnGluArgAlaGlnAlaTropJyGluArgLeuArgAlaArgMetGluMetLys 240
XX 675 CTACAGAGCGCGCGCGCGCTGGGCGAGCGGCTGCGCGCGCGATGAGAGATGGCG 734
XX
XX 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
XX 735 AGCCGACCGCGGACCGCGCTGAGCAGGTGAAGAGCAGGTGGCGAGTGGCGCCAG 794
XX
XX 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleAlaPheGlnAla 277
XX 795 CTGAGAGAGCAGCGCCGACAGATACGCTGCAGCGCGAGCGCTTCCAGCGCC 845
XX
XX RESULT 3
XX ABA83113
XX ID ABA83113 standard; DNA; 1147 BP.
XX
XX ABA83113:
XX
XX 08-FEB-2002 (first entry)
XX
XX Apolipoprotein E ovarian tumour marker gene, SEQ ID NO:63.
XX
XX DE
XX XX
XX KW Ovarian tumour marker gene; human; overexpression; upregulation;
XX KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;
XX KW identification; serous cystadenoma; borderline serous tumour;
XX KW serous cystadenocarcinoma; mucinous cystadenocarcinoma;
XX KW mucinous cystadenoma; borderline mucinous tumour; endometrioid carcinoma;
XX KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;
XX KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;
XX KW immune response pathway; cell proliferation regulation; protein folding;
XX KW membrane localized; secreted; therapeutic target; cytostatic;
XX KW gene therapy; vaccine; ds.
XX
XX OS Homo sapiens.
XX
XX XX
XX XX WO200175177-A2.
XX
XX 11-OCT-2001.
XX
XX 03-APR-2001; 2001WO-US10947.
XX
XX 03-APR-2000; 2000US-194336P.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;
XX WPI: 2001-626450/72.
XX P-PSDB: ABB50287.
XX
XX Detecting and identifying ovarian tumor, identifying increased risk for
XX developing ovarian cancer, and determining effectiveness of ovarian
XX cancer treatment, by measuring expression level of ovarian tumor marker
XX gene -
XX
XX Claim 23; Page 105-106; 140pp; English.
XX
XX The invention relates to methods for diagnosing and prognosing ovarian
XX tumours in an individual via the detection and measurement of the
XX expression of ovarian tumor marker genes (ABA83081-ABA83122, ABA83180,
XX ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179,
XX ABA83181 and ABA83183). The methods of the invention are useful for
XX detecting an ovarian tumour in a patient, for identifying an individual
```


at increased risk for developing ovarian cancer. In prognostic tests for assessing the relative severity of ovarian cancer, in tests for monitoring a patient in remission from ovarian cancer and in tests for monitoring disease status in a patient being treated for ovarian cancer. The methods can additionally be used to identify a particular tumour as being an ovarian tumour (i.e., an epithelial ovarian tumour selected from serous cystadenoma, borderline serous tumour, serous cystadenocarcinoma, mucinous cystadenoma, borderline mucinous tumour, mucinous cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma, clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner tumour. The ovarian tumour marker genes of the invention were identified using SAGE (serial analysis of gene expression) and were found to be overexpressed in a broad variety of ovarian epithelial tumour cells relative to normal ovarian epithelial cells. The marker genes are implicated in immune response pathways, in the regulation of cell proliferation and in protein folding, and many of these are membrane-localised or secreted. In addition to their use as diagnostic and prognostic markers, the ovarian tumour marker genes or their encoded proteins may be used as therapeutic targets for the treatment and prevention of ovarian cancer. Sequences ABA81081-ABA8122, ABA8180, ABA83182 and ABA83184 represent the ovarian tumour marker genes of the invention.

Sequence 1147 BP; 210 A; 365 C; 425 G; 147 T; 0 other:

Alignment Scores:

Pred. No.:	2.28e-102	Length:	1147
Score:	1374.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	22	Gaps:	0

US-09-827-854-15_COPY_1_277 (1-277) x ABA83113 (1-1147)

```

QY 1 MetIysValLeuTrpAlaIleuLeuValThrPheLeuAlaIlysgInAlaIlyVal 20
DB 46 ATGAAGTTCTGTGGGCTGCTGCTGTGTGCATTCCTCGCAGAGTCCAGGCAAGGCTG 105
QY 21 GAGGAAIAlaIgluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 106 GAGCAAGCGGTGAGACAGACAGCGGAGCCGAGCTCGCCAGACAGACCGAGTGGCAGAC 165
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 166 GGGCAGCGCTGGGAAGTGGAGCTGGCTTTGGATTACTGGCGTGGGCGAGACA 225
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 226 CTGCTGAGCAGGTGACAGAGAGGCTGCTCAGCTCCAGGTCACCCAGGAACTGAGGCGG 285
QY 81 LeuMetAspGluThrMetIysGluLeuAlaIlyTrpIysSerGluLeuGluGlnGlnLeu 100
DB 286 CTGATGAGACAGACCATGAAAGAGTTGAAGGCTTACAAATCGGAAGTGGAGACAACG 345
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerIysGluLeuGlnAlaIlyAla 120
DB 346 ACCCGCGTGGCGAGAGACAGCGGCGACGCTGCCAAGAGCTGAGCGGCGGCGAGGCC 405
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 406 CGGCTGGGCGCGACATGAGAGACGCTGCGGCGCTGTCAGTACACGCGGCGAGGCTG 465
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluValArgLeuAlaSerHisLeuArg 160
DB 466 CAGGCCATGCTGGCGAGACACCGAGAGCTGGGCGCTCCCTCCACCTGCGCG 525
QY 161 LysLeuArgIysArgLeuLeuArgAspAlaAspAspLeuGlnIlyArgLeuAlaValTyr 180
DB 526 AAGCTCGTAAAGGCGCTCTCCGCGATGCGGATGACCTCGACAGAGCGCTGGAGAGTAC 585
QY 181 GlnAlaGlyAlaArgGluGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200

```

```

DB 586 CAGCGCGGCGCGCGAGCGCGCCGAGCGGCGCTCAGCGCCATCCCGACGCGCTGGG 645
QY 201 ProLeuValGluGlnIlyArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
DB 646 CCCCTGTGGAAACAGGCGCGCGTGGCGCGCACGTGTGGGCTCCCTGGCGCGCGAGCGG 705
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
DB 706 CTACAGAGCGGCGCGCAGCGCTGGGCGAGCGGCTGCGCGCGGATGAGAGATGGCG 765
QY 241 SerArgThrArgAspArgLeuAspGluValIysGluGlnValAlaGluValArgAlaIly 260
DB 766 AGCGGAGCGCGAGCGCGCTGGAGAGTGAAGAGACAGTGGGAGGTGCGCGCGAAG 825
QY 261 LeuGluGluGlnAlaGlnIlyArgLeuGlnAlaGlnAlaPheGlnAla 277
DB 826 CTGAGAGAGCGAGCGCCAGCAGATACGCTGACGCGGAGGCGCTTCCAGGCC 876

```

RESULT 4

AAAF84315 standard; cDNA; 1156 BP.

AAAF84315;

21-JUN-2001 (first entry)

Human APOE3 coding sequence.

Human; APOE3; Alzheimer's disease; arteriosclerosis; ss.

Homo sapiens.

Key Location/Qualifiers

FT CDS 61..1014

FT /tag= a

FT /product= "human APOE3"

PN JP2001017028-A.

PD 23-JAN-2001.

PE 28-APR-2000; 2000JP-0128919.

PR 06-MAY-1999; 99JP-0125647.

PA (MITU) MITSUBISHI CHEM CORP.

DR WPI; 2001-285406/30.

DR P-PSDB; AAB80997.

PT New apoe humanized mammalian cell useful for screening for agents

PT useful for treating or preventing Alzheimer's disease and

PT arteriosclerosis -

PS Disclosure; Page 13-14; 22pp; Japanese.

XX The present invention relates to an APOE humanised mammalian cell. The

XX present sequence is the coding sequence for human APOE3, which was used

XX in the method of the present invention. The APOE humanised mammalian cell

XX can be used for screening for agents useful for treating or preventing

XX Alzheimer's disease and arteriosclerosis.

US-09-827-854-15_COPY_1_277 (1-277) x AAF84315 (1-1156)

Alignment Scores:

Pred. No.:	2.3e-102	Length:	1156
Score:	1374.00 <td>Matches:</td> <td>277</td>	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	22	Gaps:	0

QY	1	MethylValLeuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
DB	61	ATGAAGGTTCTGTGGGCTGGCTGGTGCATTCCTCGGAGAGATGCCAGGAGGTG	120
QY	21	GIUGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer	40
DB	121	GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGGCCGACAGAGAGAGTGGCAGAGC	180
QY	41	GIyGlnArgTrpGluLeuAlaLeuGlyAArgPheTrpAspTyrLeuArgTrpValGlnThr	60
DB	181	GGCCAGCGCTGGGAAGTGGCACTGGTGGCTTTTGGATACCTGGCTGGTGGTGCAGACA	240
QY	61	LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80
DB	241	CTGTCTGAGCAGGTGCAGAGAGCTGCTCAGCTCCAGGTCACCCAGAGAACTAGAGGCG	300
QY	81	LeuMetAspGluThrMetLeuGlyLeuLysAlaTyrTyrLysSerGluLeuGluGlnLeu	100
DB	301	CTGATGGACGAGACCATTAAGAGAGTGAAGCTACCAATGGAAGTGGAGAACTG	360
QY	101	ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
DB	361	ACCCGCGGTGGAGAGACACCGGCGAGCGCTGTCCAAGAGAGTGCAGGCGCCAGAGCC	420
QY	121	ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal	140
DB	421	CGGGGGGCGGAGATGAGAGAGCTGTGCGCGCCCTGGTGCAGTACCCGCGGAGAGTG	480
QY	141	GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg	160
DB	481	CAGGCGATGCTCGGCGAGACACCGAGAGAGCTGGGGTGGCTTCGCTCCACTGCGCC	540
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
DB	541	AAGCTGCGTAGCGGCTCTCCGCGATGCCATGACCTGCAGAGAGCGCTGCGAGTGAC	600
QY	181	GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200
DB	601	CAGGCGGCGGCGCGAGAGCGCGAGCGGCGCTCAGCCGCAACCGGAGCGCTGGGG	660
QY	201	ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGluPro	220
DB	661	CCCTGCTGGGAGACAGGGCGCGCTGGGCGCCCACTGTGGGCTCTCTGGCGGCGAGCG	720
QY	221	LeuGlnGluArgAlaGlnAlaTrpGlyLysArgLeuArgAlaArgMetGluMetGly	240
DB	721	CTACAGAGAGCGGGCCAGAGCTGGGGGAGCGGCTGGCGGCGGAGATGAGATGGGC	780
QY	241	SerArgThrArgAspArgLeuAspGluValLysGluGlnGlnAlaGluValArgAlaLys	260
DB	781	AGCCGAGACCCCGACCGCTGTGAGAGTGAAGAGAGCAGGTGGCGGAGTGGCGCCAG	840
QY	261	LeuGlnGluGlnAlaGlnGlnGlnIleArgLeuGlnGlnAlaGlnAlaPheGlnAla	277
DB	841	CTGGAGAGAGCGCCACAGATACGCTGCGAGGCCGAGGCGCTTCCAGGCC	891
RESULT 5	AAD22048		
ID	AAD22048	standard; DNA: 1156 BP.	
XX	AC		
XX	AD22048;		
XX	12-FEB-2002	(first entry)	
XX	DT		
XX	DE		
XX	Human apolipoprotein E (apoE)	isoprotein, apoE3 DNA.	
XX	Human: apolipoprotein E; apoE; cholesterol: atherosclerosis;		
XX	hypertriglyceridaemia; low density lipoprotein; LDL; ds.		
XX	Homo sapiens.		
XX	Key	Location/Qualifiers	
XX			

Claim 14; Page 81; 91pp; English

The present sequence is a human apolipoprotein E (apoE) isoform, apoE3 DNA. The apoE lipoproteins are useful for lowering cholesterol, delaying the onset of atherosclerosis, treating or preventing atherosclerosis without inducing hypertriglyceridaemia, in a mammal lacking an endogenous, normally functioning apoE gene or low density lipoprotein (LDL) receptor or is at risk for developing atherosclerosis due to accumulation of lipoprotein remnants in the bloodstream or having a defect in remnant removal.

Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other;

ment Scores:

No.:	2,3e-102	length:	1156
Matches:	1374.00	Conservative:	277
ent Similarity:	100.00%	Mismatches:	0
Local Similarity:	100.00%	Indels:	0
y Match:	100.00%	Gaps:	0
	24		

9-827-854-15_COPY_1_277 (1-277) x AAD22048 (1-1156,

1 MetLysValIleuTriPalaIalaIeuIeuVal1ThrPheIeuIalagIcysGlnIalalyVal 20
61 ATGAAAGTTCCTGGGCTCCGTTGCTGGTCACATTCCTGGAGGATCCAGGCCAAAGTG 120
21 GlnGlnIalavalGIuThrGluProGluProGluIeuArgIuGlnIuThrGluTrpGlnSer 40
121 GAGCAAGCCGCTGGAAACAAGGCCGGAAGCCGACGCTGGCCCAAGACCGAGTGGCAAGC 180
41 GlnGlnArgTrpGluIeuIalalauIeuIalylarPheTrpAspTrpIleuArgTrpValGIuThr 60
181 GGCACACGCTGGGAAGCTGGCACTGGGCTGGCTTTGGGATTCCTACGCTGGTGTCAACA 240
61 LeuSerGlnGlnValGlnGlnIuIeuIeuSerSerGlnVal1ThrGlnIuIeuArgAla 80
241 CTCTCGAACACAGTCCAGAGAGAGCTGCTCAGCTCCAGGTACACCCAGAACTGAGGGCG 300
81 LeuMetAspGluThrMetLysGluIeuLysAlaTrpLysSerGluIeuGluGluGlnIeu 1000


```
Db 301 CTATGAGACAGACCATGAAGGAGTTGAAGCCTACAAATCGGAAGGAGCAACTG 360
Qy 101 ThProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 361 ACCCGGTGGCGGAGAGACGGGGCCAGGCTGTCCAAAGAGAGTGCAGGGCGGCGAGGCC 420
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 421 CGGCTGGGCGGACATGAGGAGAGCTGTGGCGCGCTGTGTGACGTACCCGGCGAGAGTG 480
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCATGCTCGGCCACAGACACCGAGAGCTCGGGTGGCGCTCCGCTCCACCTGCC 540
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AACCTGCGTAAGCGGCTCTCCCGCATGCCGATGACCTGCAGAAAGCGCCTGGCAGTGTAC 600
Qy 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 601 CAGGCGGGGCGCGCGAGGGCGCGGAGCGCGGCTCAGCCCATCCGCGAGCGCTGGGG 660
Qy 201 ProLeuValGluGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGTGACACAGAGGCGCGCTGTGGCGCGCCACTGTGGGCTCCTGGCGGCGACCG 720
Qy 221 LeuGlnGluArgAlaGlnAlaTyrGlyGluArgLeuArgAlaArgMetGluMetGly 240
Db 721 CTACAGAGAGGCGGCGCGGCTGTGGGGAGCGGCTGGCGCGGAGTGCAGAGATGGCG 780
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 781 AGCGGAGCGCGCGACGCGCTCGAGAGGTGAAGAGAGAGTGGCGGAGGTGGCGCGCAAG 840
Qy 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaGlnAla 277
Db 841 CTGAGAGAGAGCGCCAGCATACGCTTCGAGGCGAGGCGCTTCAGGCC 891
RESULT 6
AADD2052
ID AADD2052 standard; DNA; 1156 BP.
XX AC AADD2052;
XX 12-FEB-2002 (first entry)
XX DE Human apolipoprotein E (apoE) allele, apoE2** DNA.
XX KW Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;
XX hypertriglyceridaemia; low density lipoprotein; LDL; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 61..1014
XX FT sig_peptide /product= "Human apoE allele, apoE2**"
XX FT mat_peptide /tag= b
XX FT /tag= 115..1011
XX FT /product= "Mature human apoE allele, apoE2**"
XX PN WO200177136-A1.
XX PD 18-OCT-2001.
XX PF 06-APR-2001; 2001WO-US1358.
XX PR 06-APR-2000; 2000US-0544386.
XX PR 04-OCT-2000; 2000US-0679088.
XX PR 05-APR-2001; 2001US-0827854.
XX
```

```
PA (KOSP-) KOS PHARM INC.
PA (UYBO-) UNITV BOSTON.
XX
PI Zannis VI, Kypros KE;
XX
DR WPI: 2002-010885/01.
DR P-PDB: AAEL3298.
XX
PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT cholesterol, delaying the onset of or treating atherosclerosis in
PT mammal, without inducing hypertriglyceridaemia
XX
PS Claim 14; page 83; 91pp; English.
XX
CC The present sequence is a human apolipoprotein E (apoE)
CC allele, apoE2** DNA. The apoE lipoproteins are useful for
CC lowering cholesterol, delaying the onset of atherosclerosis,
CC treating or regressing atherosclerosis without inducing
CC hypertriglyceridaemia, in a mammal lacking an endogenous,
CC normally functioning apoE gene or low density lipoprotein (LDL)
CC receptor or is at risk for developing atherosclerosis due to
CC accumulation of lipoprotein remnants in the bloodstream or having
CC a defect in remnant removal.
XX
SQ Sequence 1156 BP; 207 A; 369 C; 432 G; 148 T; 0 other:
Alignment Scores:
Pred. No.: 4,84e-102 Length: 1156
Score: 1370.00 Matches: 276
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.64% Mismatches: 0
Query Match: 99.71% Indels: 0
DB: 24 Gaps: 0
US-09-827-854-15_copy_1_277 (1-277) x AADD2052 (1-1156)
Qy 1 MetLysValLeuThrPalaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGTTCTGTGGGCTGCTGTGTGTCATTCCTCGCAGAGATCCAGGCCAAGGTG 120
Qy 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnIleThrGluThrPglIns 40
Db 121 GAGCAAGCGGTGGAGACAGAGCGGAGCGCGGACCTCGCGCAGACGCGAGTGGCAGAGC 180
Qy 41 GlyIleArgTrpGluLeuAlaLeuGlyArgPheThrAspTyrLeuArgTrpValGlnThr 60
Db 181 GGGCAGCGCTGGGAACTGGACCTGGGCTTTGGGATTACTGGGCTGGTGGCAGACA 240
Qy 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValIleThrGlnLeuArgAla 80
Db 241 CTGCTGAGCAGCTGCAGAGAGAGTGTACCTCCAGGTACCCAGCAACTGAGGGCG 300
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
Db 301 CTGATGAGACAGACCATGAAGGAGTTGAAGCCTACAAATCGGAAGGAGCAACTG 360
Qy 101 ThProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 361 ACCCGGTGGCGGAGAGACGGGGCCAGGCTGTCCAAAGAGAGTGCAGGGCGGCGAGGCC 420
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 421 CGGCTGGGCGGACATGAGGAGAGCTGTGGCGCGCTGTGTGACGTACCCGGCGAGAGTG 480
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCATGCTCGGCCACAGACACCGAGAGCTCGGGTGGCGCTCCGCTCCACCTGCC 540
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AACCTGCGTAAGCGGCTCTCCCGCATGCCGATGACCTGCAGAAAGCGCCTGGCAGTGTAC 600
Qy 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
```



```

|||||
Db 601 CAGCGCGGGGCCCGCAGGGCCGCGGCTCAGCGCCATCCGCGCGCCCTGGGG 660
QY 201 ProLeuValGluGlnGlyArgValArgAlaIleThrValGlySerLeuAlaGlyInPro 220
Db 661 CCCCTGTGTGAACAGAGCGCGCGTGGCGCCGACACTGTGTGGCTCCCTGGCCGCGCCG 720
QY 221 LeuGlnGluArgAlaGlnAlaIleThrPglIyGluArgLeuArgAlaArgMetGluGluMetGly 240
Db 721 CTACAGAGAGCGGGCCAGAGCTGGGGCGAGCGGCTGGCGCGGATGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 781 AGCCGAGACCCGCGACCGCTGGCGAGGTGAAGAGCAGGTGGCGGAGGTGGCGCCAG 840
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAla 277
Db 841 CTGGAGAGCAGCGCCAGAGATACGCTTCAGAGCGCGGCTTCAGAGCC 891

```

RESULT 7
AAF84314
ID AAF84314 standard; cDNA; 1156 BP.

```

XX AC AAF84314;
XX XX
XX 21-JUN-2001 (first entry)
XX DE Human ApoE2 coding sequence.
XX XX
XX Human; ApoE2; Alzheimer's disease; arteriosclerosis; ss.
XX OS Homo sapiens.
XX XX
XX FH Key Location/Qualifiers
XX FT CDS 61..1014
XX FT /tag=a
XX FT /product="Human ApoE2"
XX PN JP2001017028-A.
XX PD 23-JAN-2001.
XX PF 28-APR-2000; 2000JP-0128919.
XX PR 06-MAY-1999; 99JP-0125647.
XX PA (MITU ) MITSUBISHI CHEM CORP.
XX DR WPI: 2001-285406/30.
XX DR P-PSDB; AAB80996.
XX XX
XX PT New apoe humanized mammalian cell useful for screening for agents
XX PT useful for treating or preventing Alzheimer's disease and
XX PT arteriosclerosis -
XX PS Disclosure: Page 11-12; 22pp; Japanese.
XX XX
XX CC The present invention relates to an ApoE humanised mammalian cell. The
XX CC present sequence is the coding sequence for human ApoE2, which was used
XX CC in the method of the present invention. The ApoE humanised mammalian cell
XX CC can be used for screening for agents useful for treating or preventing
XX CC Alzheimer's disease and arteriosclerosis.
XX XX
XX SQ Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;

```

Alignment Scores:

```

Pred. No.: 1.02e-101
Score: 1366.00
Percent Similarity: 99.64%
Best Local Similarity: 99.64%
Query Match: 99.42%
DB: 22
Gaps: 0

```

us-09-827-854-15_copy_1_277 (1-277) x AAF84314 (1-1156)

```

QY 1 MetLysValLeuThrPheAlaLeuLeuValThrPheLeuAlaGlyGlnAlaLysVal 20
Db 61 ATGAAAGTTCTGTGGGCTGCTGCTGTGTGACATTCCTGGCAGAGGCCAAGGAGTG 120
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrPglIns 40
Db 121 GAGCAAGCGGTGAGAGACAGACCGGAGCCGAGCTGCGCAGCAGACCCAGTGGCAGAGC 180
QY 41 GlyLysArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
Db 181 GGCACACCGCTGGGAACATGACACTGGGTCCCTTTGGATACCTGCGCTGGGTGCAGACA 240
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGTCTGACAGAGTGAGAGAGAGCTGCTCAGCTCCAGATCCAGAGAACTGAGAGGGC 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnGlnLeu 100
Db 301 CTGATGACAGACACCATGAAGAGTTGAAGCCCTTACAAATCGGAACCTGGAGAACACTG 360
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 361 ACCCGGTGGCGAGAGAGAGAGCGGCGCAGCGCTGTCCAAAGAGCTGCAGCGCGCGAGGC 420
QY 121 ArgLeuGlyAlaAspMetLysAspValLysGlyArgLeuValGlnTrpArgGlyGluVal 140
Db 421 CGGCTGGGCGGACATGAGAGAGCTGTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTGCGCCAGAGCAGCAGCAGAGCTGCGGGTGCCTGCCCTCCACCTGGCC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTCGTAAAGCGGCTCTCCGCGATGCGCATGACCTGACAGAAAGTGCCTGCGAGTAC 600
QY 181 GlnAlaGlyAlaArgGluGluLysAlaGluArgGlyLeuSerAlaIleArgGluThrArgLeuGly 200
Db 601 CAGGCGGGGGCCCGCAGAGGCGCGCGGCTCAGCGCCATCCGCGAGCGCCCTGGGG 660
QY 201 ProLeuValGluGlnGlyArgValArgAlaIleThrValGlySerLeuAlaGlyInPro 220
Db 661 CCCCTGTGTGAACAGAGCGCGCGTGGCGCCGACACTGTGTGGCTCCCTGGCCGCGCCG 720
QY 221 LeuGlnGluArgAlaGlnAlaIleThrPglIyGluArgLeuArgAlaArgMetGluGluMetGly 240
Db 721 CTACAGAGAGCGGGCCAGAGCTGGGGCGAGCGGCTTCGCGCGGATGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 781 AGCCGAGACCCGCGACCGCTGGCGAGGTGAAGAGCAGGTGGCGGAGGTGGCGCCAG 840
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAla 277
Db 841 CTGGAGAGCAGCGCCAGAGATACGCTTCAGAGCGCGGCTTCAGAGCC 891

```

RESULT 8

AAD22049
ID AAD22049 standard; DNA; 1156 BP.

XX AAD22049;

XX 12-FEB-2002 (first entry)

XX Human apolipoprotein E (apoE) isoprotein, apoE2 DNA.

XX Human; apolipoprotein E; apoE; cholesterol; arteriosclerosis;

XX hypertriglyceridaemia; low density lipoprotein; LDL; ds.

XX Homo sapiens.

FH	Key	Location/Qualifiers
FT	CDS	61..1014
FT		/*tag= a
FT	sig_peptide	/product= "Human apoE isoform, apoE2"
FT		61..114
FT		/*tag= b
FT	mat_peptide	115..1011
FT		/*tag= c
FT		/product= "Mature human apoE isoform, apoE2"
XX		
PN	MO200177136-A1.	
PD	18-OCT-2001.	
XX		
PF	06-APR-2001; 2001WO-US11358.	
XX		
PR	06-APR-2000; 2000US-0544386.	
PR	04-OCT-2000; 2000US-0679088.	
PR	05-APR-2001; 2001US-0827854.	
XX		
PA	(KOSP-) KOS PHARM INC.	
PA	(UYBO-) UNIV BOSTON.	
P1	Zannis VI, Kypros KE;	
XX		
DR	WPI: 2002-010885/01.	
DR	P-PSDB; AAELJ295.	
PT	New apolipoprotein E polypeptide and nucleic acid, useful for lowering	
PT	cholesterol, delaying the onset of or treating atherosclerosis in	
PT	mammals, without inducing hypertriglyceridemia	
XX		
PS	Claim 14: Page 81-82; 91pp; English.	
CC	The present sequence is a human apolipoprotein E (apoE)	
CC	isoform, apoE2 DNA. The apoE lipoproteins are useful for	
CC	lowering cholesterol, delaying the onset of atherosclerosis,	
CC	treating or regressing atherosclerosis without inducing	
CC	hypertriglyceridemia, in a mammal lacking an endogenous,	
CC	normally functioning apoE gene or low density lipoprotein (LDL)	
CC	receptor or is at risk for developing atherosclerosis due to	
CC	accumulation of lipoprotein remnants in the bloodstream or having	
CC	a defect in remnant removal.	
XX		
SQ	Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;	
	Alignment Scores:	
	Pred. No.: 1,02e-101 Length: 1156	
	Score: 1366.00 Matches: 276	
	Percent Similarity: 99.64% Conservative: 0	
	Best Local Similarity: 99.64% Mismatches: 1	
	Query Match: 99.42% Indels: 0	
	Gaps: 24	
US-09-827-854-15_COPY_1_277 (1-277) x AAD22049 (1-1156)		
OY	1 MetlysvallenturpalalaaleulevalthrPhelenuaglYCyslnalalyval	20
Dd	61 ATGAAGGTTTCGTGGCGTGGTGTCGACATCTCCTGCAGAGATGCCAAGGTG	120
OY	21 GlutlnlAlavalgluthrGluprogluproluleuArnglnglnthrlutrpqlnsr	40
Dd	121 GAGCAACCGGTGGAGACAGACCAGCCGACACTCGCACAGACAGCCAGTGGCAGAGC	180
OY	41 GlycInlrgrTPgIuleuAlaleuGLyArgPheTrAspTyrlLeuArgrtPrvAlcInthr	60
Dd	181 GCCAGCGCTGGAGACTGGCAGCTGGGCGCTTTGGGATTACCTGGCGGTGGCAGACA	240
OY	61 LeusergluInValGlnGlnGluInLeuLeuSerSerglnValThrGlnGluLeuArrgla	80
Dd	241 CTGCTGAGCAGGTGGCAGAGAGAGAGCTGCTACGCTCCAGGTACCCAGAACTAGGGCG	300
OY	81 LeuMetAspgIutrhMetlysgIuleuLyAlatyrLySercluleuGlnGlnleu	100

Db	301	CTGATGGACGAGACCATTAAGAGAGTTGAAGGCTTACCAATGTGGAACTGGAGAACCAACTG	360
Qy	101	ThrProValAlaGluGluThrArgAlaArgLeuSerIytsGluLeuGlnAlaAlaGlnAla	120
Db	361	ACCCGGGTGGCGGAGAGACGCGGGGCGACGGCTGTCTCAAGAGCTGCAGAGCGCGCGCACGGCC	420
Qy	121	ArgLeuGlyAlaAspMetGluAspValCysGlyIarGluValGlnIYrArGlyGluVal	140
Db	421	CGCGTGGCGCGGACATGAGAGACCTGTGCGCGCGCCCTGGGTGCAGTACC CGCGGAGAGT	480
Qy	141	GlnAlaMetLeuGlyGlnSerThrGluIuLeuNyrGValArgLeuAlaIaserHisLeuArg	160
Db	481	CMGGCCATGCTGGCGCAAGCAGCAGAGAGCTGCGGGTGCCTCTCGCCTCCACCTGCGC	540
Qy	161	LysLeuArgIySarLeuLeuAlaArgAspAlaAspAspLeuGlnIytsArgLeuAlaValTYr	180
Db	541	AAGCTGCGGTAAAGGGCTCTCTCCGGATGCGATGACCTGCAGAGAGTGCCTGGCAGGTAC	600
Qy	181	GlnAlaGlyAlaArgGluGlyAlaGluArgIyLeuSerAlaIleArgGluArgLeuGly	200
Db	601	CMGGCGGGGCGCGGAGAGGCGCGAGCGCGCTTCAGACCGCAATCCCGAGCGCTTGGG	660
Qy	201	ProLeuValGluGlnGlyArgValArgAlaIarThrValGlySerLeuAlaGlyGlnPro	220
Db	661	CCCCGTGTGGACACAGGCGCGGCTGCGGGCGCCCACTGTGGGCTCTCCGTGGCGGCGACCG	720
Qy	221	LeuGlnGluArgAlaGlnAlaATrPolYcIuArgLeuArgAlaArMetGluGluMetGly	240
Db	721	CTACAGAGAGGCGGCGCCACAGGCTCGGGGCGCAGCGCGTGGCGCGCGGAGTGAAGATGGC	780
Qy	241	SerArgThrArGAspArgLeuAspGluValIytsGluGlnValAlaGluValArgAlaIyts	260
Db	781	AGCCCGACCCCGGACCGGCTGTGAGAGAGGTGAAGAGCAGGTGGCGGAGGTGCGCGCCACG	840
Qy	261	LeuGlnGluGlnAlaGlnGlnIleArGleuGlnIuArgIuAlaPheGlnAla	277
Db	841	CTGGAGAGAGAGGCGCCAGACAGATACGCTGCGAGCGCCGAGGCTTCCAGGCC	891
RESULT 9			
AAAD22051	ID	AAAD22051 standard; DNA; 1156 BP.	
XX	AAAD22051;		
AC			
XX			
DT	12-FEB-2002 (first entry)		
XX			
DE	Human apolipoprotein E (apoE) allele, apoE2* DNA.		
XX			
KM	Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;		
KM	hypertriglyceridaemia; low density lipoprotein; LDL; ds.		
XX			
OS	Homo sapiens.		
XX			
FT	Key	Location/Qualifiers	
FT	CDS	61..1014	
FT		/*tag= a	
FT		/*product= "Human apoE allele, apoE2*"	
FT		61..114	
FT		/*tag= b	
FT		115..1011	
FT	mat_peptide	/*tag= c	
FT		/*product= "Mature human apoE allele, apoE2*"	
XX			
PN	WO20017136-A1.		
XX			
PD	18-OCT-2001.		
XX			
PE	06-APR-2001; 2001WO-US11358.		
XX			
PR	06-APR-2000; 2000US-0544386.		
PR	04-OCT-2000; 2000US-0679088.		
PR	05-APR-2001; 2001US-0837854.		

CC used in therapeutic applications.

XX Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;

Alignment Scores:

Pred. No.:	1,23e-101	Length:	1157
Score:	1365.00	Matches:	275
Percent Similarity:	99.28%	Conservative:	0
Best Local Similarity:	99.28%	Mismatches:	2
Query Match:	99.34%	Indels:	0
DB:	17	Gaps:	0

US-09-827-854-15_COPY_1_277 (1-277) x AAT06957 (1-1157)

```

QY 1 MetlyValLeuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
   |||||
Db 62 ATGAGGTTCTGTGGCTGGTGTGCTGTCACATTCCTGGCAGAGATCCAGGCCAAGT 121
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
   |||||
Db 122 GAGCAAGCGGTGAGACAGACCGGAGCCGAGCTCGCGCAGACAGACCGAGTGGCAGAGC 181
QY 41 GlyAlaArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
   |||||
Db 182 GGCACAGCGGTGGAACTGGCACTGGGTGCTTTGGATTACCTGGCTGGGTGCAGACA 241
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
   |||||
Db 242 CTGCTGAGCAGAGGTGAGGAGAGAGCTCTCAGCTCCCAATCCACCAAGAACTGAGGCG 301
QY 81 LeuMetAspGluThrMetTyrGluLeuValTyrTyrSerGluLeuGlnGlnLeu 100
   |||||
Db 302 CTGATGACAGACCAATGAGGAGTTGAAAGCTTACAAATCGAAGGAGGAACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerTyrGluGlnAlaGlnAla 120
   |||||
Db 362 ACCCGGTTAGCGAGAGAGACCGGGCGACGCTGTCCAGAGACTGCGAGCGCGCAGGCC 421
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
   |||||
Db 422 CGCTGGCGCGGACATGAGAGACGTGTGGCGCGCTGTGCAGTACCCGCGGAGAGT 481
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
   |||||
Db 482 CAGGCGCATGCTCGCCAGACACCGAGAGCTCGGGTGGCGCTCCGCTCCACCTGCGC 541
QY 161 LysLeuArgTyrArgLeuLeuArgAspAlaAspAspLeuGlnTyrArgLeuAlaValTyr 180
   |||||
Db 542 AAGCTGGGTAAAGCGGCTCTCCCGGATCCCGATGACTCTCAGAAAGCCCTGGCAGTAC 601
QY 181 GlnAlaGlyAlaArgGluGlnAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
   |||||
Db 602 CAGCGCGGCGCGCGGAGCGCGGAGCGCGGCTCAGCCATCCGCGAGCGCGCTGGG 661
QY 201 ProLeuValGluGlnGlnTyrArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
   |||||
Db 662 CCCCTGGTGAACAGGCGCGCGGCGCGGCGCAGTGTGGCTCCCTGGCGCGCGCAGCCG 721
QY 221 LeuGlnGluArgAlaGlnAlaTyrGlyArgLeuArgAlaArgMetGluGlnMetGly 240
   |||||
Db 722 CTACAGAGAGCGGCGCGGCTGGGCGAGCGGCTGGCGCGCGGATGAGGAGATGGCG 781
QY 241 SerArgThrArgAspArgLeuAspGluValTyrGlnGlnValAlaGluValArgAlaTyr 260
   |||||
Db 782 AGTCGAGCGCGCGCGCGCTGGAGCAGAGTGAAGAGAGAGTGGCGGAGCGTGGCGCGCAG 841
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 277
   |||||
Db 842 CTGAGGAGAGCGCGCGCAGCAGATACGCTGCGAGCGAGCGCTTCCAGGCC 892

```

RESULT 11
ABN95746
ID ABN95746 standard; DNA: 1157 BP.

```

XX AC ABN95746;
XX XX 13-AUG-2002 (first entry)
DT DT
XX DE Gene #2244 used to diagnose liver cancer.
XX DE
XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
XX KW metastatic liver tumour; cytostatic; expression profile; disease state;
XX KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX OS Homo sapiens.
XX PN WO200229103-A2.
XX PD 11-APR-2002.
XX PF 02-OCT-2001; 2001WO-US30589.
XX PR 02-OCT-2000; 2000US-237054P.
XX PA (GENE-) GENE LOGIC INC.
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX WPI: 2002-426119/45.
XX DR
XX PT diagnosing and detecting the progression of liver cancer,
XX PT hepatocellular carcinoma or metastatic liver tumor in a patient,
XX PT involves detecting the level of expression of two or more genes in a
XX PT liver tissue sample
XX PS
XX XS Claim 1: SEQ ID NO 2244; 298bp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytosolic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;

```

Alignment Scores:

Pred. No.:	1,23e-101	Length:	1157
Score:	1365.00	Matches:	275
Percent Similarity:	99.28%	Conservative:	0
Best Local Similarity:	99.28%	Mismatches:	2
Query Match:	99.34%	Indels:	0
DB:	24	Gaps:	0

US-09-827-854-15_COPY_1_277 (1-277) x ABN95746 (1-1157)

```

QY 1 MetlyValLeuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
   |||||
Db 62 ATGAGGTTCTGTGGCTGGTGTGCTGTCACATTCCTGGCAGAGATCCAGGCCAAGT 121
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
   |||||
Db 122 GAGCAAGCGGTGAGACAGACCGGAGCCGAGCTCGCGCAGACAGACCGAGTGGCAGAGC 181
QY 41 GlyAlaArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
   |||||
Db 182 GGCACAGCGGTGGAACTGGCACTGGGTGCTTTGGATTACCTGGCTGGGTGCAGACA 241

```



```

DB 482 CAGGCCATGCTCGGCCAGACACCGAGACTCGGGTGGCTTCCCTCCCACTGCGC 541
OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 542 AAGCTGCGTAAAGGGGCTCTCCCGATCCGATGACATGACAGAAAGCGCTGGCAGTGTAC 601
OY 181 GlnAlaGlyAlaArgGluGluGluAlaValArgGlyLeuSerAlaIleArgGluValGluGly 200
DB 602 CAGGCGGGGGCCCGCAGAGGGCGCGAGCGCGCTCAGCCCTCCGCGCGCGCGCGGG 661
OY 201 ProLeuValGluGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 662 CCCCCTGGTGAACAGGGCCCGCTGCGGGCGCCCACTGTGGGCTCCCTGGCGGGCCAGCG 721
OY 221 LeuGlnGluArgAlaGlnAlaIleArgTpsGlyGluArgLeuArgAlaArgMetGluMetGly 240
DB 722 CTACAGAGAGCGGCGCCAGGCGCTGGGCGAGCGCGCTGCGCGCGCGGATGAGAGATGGCG 781
OY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 782 AGTCGAGACCGCGCAGCCCGCTGGAGAGGTGAAGAGCAGGTGGCGGAGGTGCCGCAAG 841
OY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleArgGlnAla 277
DB 842 CTGAGAGAGACGAGCCACGACGATACGCTGCGAGCGCGCGCTTCCAGAGCC 892

RESULT 13
ABL65450
ID ABL65450 standard; DNA; 1157 BP.
XX
AC ABL65450;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:3787.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001MO-US10838.
XX
XX 05-JUN-2000; 2000US-209473P.
XX 05-JUN-2000; 2000US-209531P.
XX 18-SEP-2000; 2000US-233133P.
XX 18-SEP-2000; 2000US-233617P.
XX 20-SEP-2000; 2000US-234009P.
XX 20-SEP-2000; 2000US-234034P.
XX 20-SEP-2000; 2000US-234052P.
XX 22-SEP-2000; 2000US-234509P.
XX 22-SEP-2000; 2000US-234567P.
XX 25-SEP-2000; 2000US-234923P.
XX 25-SEP-2000; 2000US-234934P.
XX 25-SEP-2000; 2000US-235077P.
XX 25-SEP-2000; 2000US-235082P.
XX 25-SEP-2000; 2000US-235134P.
XX 26-SEP-2000; 2000US-235280P.
XX 26-SEP-2000; 2000US-235637P.
XX 26-SEP-2000; 2000US-235638P.
XX 27-SEP-2000; 2000US-235711P.
XX 27-SEP-2000; 2000US-235720P.
XX 27-SEP-2000; 2000US-235840P.
XX 27-SEP-2000; 2000US-235863P.
XX 28-SEP-2000; 2000US-236028P.
XX 28-SEP-2000; 2000US-236032P.

```

```

PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.

PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrikan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
XX PT chemical agent to be tested for anti-neoplastic activity, and
XX PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1; SEQ ID 3787; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an
XX CC anti-neoplastic agent. The method involves exposing cells to a chemical
XX CC agent to be tested for anti-neoplastic activity, determining a change in
XX CC expression of at least one gene (I) of a signature gene set, where (I)
XX CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX CC to ABL70110), or is at least 95% identical to (S), where a change in
XX CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX CC activity and can be used in gene therapy. M1 can be used for screening
XX CC an anti-neoplastic agent, and can be used for producing a product which
XX CC is the data collected with respect to the anti-neoplastic agent as a
XX CC result of M1, and the data is sufficient to convey the chemical
XX CC structure and/or properties of the agent. M1 can be used in the
XX CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
XX CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
XX CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
XX CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
XX CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;
XX
XX
Alignment Scores:
Pred. No.: 1,23e-101 Length: 1157
Score: 1365.00 Matches: 275
Percent Similarity: 99.28% Conservative: 0
Best Local Similarity: 99.28% Mismatches: 2
Query Match: 99.34% Indels: 0
DB: 24 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x ABL65450 (1-1157)
OY 1 MetLysValIleuTrpAlaIleuLeuValTrpPheLeuAlaGlyCysGlnAlaLysVal 20
DB 62 ATCAAGCTTCTGTGGGCTGCTGCTGTCGTCACATCTCTGCGAGATGCCAGGCAAGG 121
OY 21 GluGlnAlaValGluThrLupProGluProGluLeuArgGlnGlnThrLupTrpGlnSer 40
DB 122 GAGCAAGCGGTGAGACACAGAGCGGAGCCGAGCTGCGCACAGACCAGTGGCGAGAC 181
OY 41 GlyGlnArgTrpGluLeuValLeuGluArgPheTrpAspTyrLeuArgTrpValGlnTrp 60

```


Db 182 GGCACACGCTGGGAACCTGCACCTGGCTTTTGGATTACCTGCTGGGTGCAGACA 241
QY LeuSerGIuGIuValGIuGIuLeuSerSerGIuValThrGIuGIuLeuArgAla 80
Db 242 CTGTCTGACGACGAGTGGAGAGGCTGCTCAGCTCCCAAGTACCAACAACTGAGGGCG 301
QY LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGIuLeu 100
Db 302 CTGATGGACGACACCATGAGAGAGTTGAAGGCTTAAATCGAATCGAAGTGGAGAACACTG 361
QY 101 ThrProValAlaGIuGIuThrArgAlaArgLeuSerLysGIuLeuGIuAlaGIuAla 120
Db 362 ACCCGGTGACGAGAGACGCGGACGCGCTGCTCAAGAGACTGCAGACGCGCAGGCC 421
QY 121 ArgLeuGIuAlaAspMetGIuAspValLysGIuArgLeuValGIuThrArgGIuVal 140
Db 422 CGGCTGGGGGCGGACATGAGAGAGCTGTGGCGCCCTGTGCTACATCCGCGGCAAGGTG 481
QY 141 GlnAlaMetLeuGIuGIuSerThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCGCATGCTCGGCGACAGACACGAGAGCTGCGGGTGCCTGCGCTCCCACTGGCGC 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGIuLysArgLeuAlaValTyr 180
Db 542 AAGCTGCGTAAACGGGCTCTCCGCGATGCCGATGACCTGCAGAAAGCGCTGGCAGGTGAC 601
QY 181 GlnAlaGIuAlaArgGIuGIuValArgGIuLeuSerAlaIleArgGIuArgLeuGIu 200
Db 602 CAGGCGGGGGGCGGAGGCGCGGCGGCGCTGCAGCGCATCCGCGAGCGCTGGGGG 661
QY 201 ProLeuValGIuGIuGIuArgValArgValArgAlaThrValGIuSerLeuAlaGIuGIuPro 220
Db 662 CCCGTGGTGGAAACAGAGGCGCGTGGGGCGCGCACTGTGGGCTCCCTGGCGGCAAGCGC 721
QY 221 LeuGIuGIuArgAlaGIuAlaThrValArgLeuArgAlaArgMetGIuGIuMetGIu 240
Db 722 CTACAGAGGCGGCGGCGGCGGCGGCGGCTGCGCGCGGATGGAGAGATGGGCG 781
QY 241 SerArgThrArgAspArgLeuAspGIuValLysGIuGIuValAlaGIuValArgAlaLys 260
Db 782 AATCGGACCGCGGACCGCTGGACGAGGTGAAGAGAGAGGTGGCGGAGGTGGCGCCCAAG 841
QY 261 LeuGIuGIuGIuAlaGIuGIuIleArgLeuGIuAlaGIuAlaPheGIuAla 277
Db 842 CTGAGGAGAGCAGGCCAGAGATAGGCTGCGAGGCGGAGGCGCTTCACAGGCC 892

RESULT 14
AAN50450
ID AAN50450 standard; DNA; 1110 BP.
XX AAN50450;
AC AAN50450;
XX
DT 09-JAN-1992 (first entry)
XX
DE Sequence encoding human apolipoprotein E.
XX
KW Hyperlipaemia; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 15..968
FT mat_peptide 69..965
FT /*tag= a
FT /*tag= b
XX JP60118189-A.
XX
XX 25-JUN-1985.
XX
XX 29-NOV-1983; 83JP-0224980.
XX
XX 29-NOV-1983; 83JP-0224980.

XX
PA (MITU) MITSUBISHI CHEM IND KK.
XX
DR WPI: 1985-188003/31.
DR P-PSDB: AAP51204.
XX
PT DNA fragment - contg. DNA which codes human apolipoprotein E for
pr treatment of hyperlipaemia.
XX
PS Claim 3; Page 484; 8pp; Japanese.
CC The sequence may be used to produce the apolipoprotein E, useful in the
CC treatment of hyperlipaemia.
XX
SQ Sequence 1110 BP; 198 A; 355 C; 414 G; 143 T; 0 other.
Alignment Scores:
Pred. No.: 1,41e-101 Length: 1110
Score: 1364.00 Matches: 275
Percent Similarity: 99.28% Conservative: 0
Best Local Similarity: 99.28% Mismatches: 2
Query Match: 99.27% Indels: 0
DB: 6 Gaps: 0
US-09-827-854-15_copy_1_277 (1-277) x AAN50450 (1-1110)
QY 1 MetLysValIleuThrAlaIleuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 15 AAGAAGGTTCTGTGGGCTGCTGTGCTGACATTCCTGGCAGATGCCAGGCAAGGTG 74
QY 21 GlnGlnAlaValGIuThrGIuProGIuLeuArgGIuGIuThrGIuTrpGIuIns 40
Db 75 GAGCAAGCGGTGGAGACACAGAGCGGAGCCCGAGCTGCCACAGAACCGAGGCGAGAC 134
QY 41 GlyIuArgTrpGIuLeuAlaLeuGIuArgPheTrpAspTyrLeuArgTrpValGIuThr 60
Db 135 GGCAGGCGCTGGGAACCTGCACATGCGCTTTGGATTACCTGCGCTGGGTGCAGACA 194
QY 61 LeuSerGIuGIuValGIuGIuLeuLeuSerSerGIuValThrGIuGIuLeuArgAla 80
Db 195 CTGTGTGAGCACGTGCAGAGAGGCTGCTCAGCTCCAGGTACCCAGAACTGAGGCGC 254
QY 81 LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGIuLeu 100
Db 255 CTGATGACGACAGACCATGAGAGAGTTGAAGGCTTAAATCGAATCGAAGTGGAGAACACTG 314
QY 101 ThrProValAlaGIuGIuThrArgAlaArgLeuSerLysGIuLeuGIuAlaGIuAla 120
Db 315 ACCCGGTGGCGGAGAGACGCGGCGGCGGCTGCTCAAGAGACTGCAGGCGCGCAAGGCC 374
QY 121 ArgLeuGIuAlaAspMetGIuAspValLysGIuArgLeuValGIuThrArgGIuVal 140
Db 375 CGGCTGGGCGGACATGAGAGAGCTGTGGCGCGCTGTGCTACATCCGCGGCAAGGTG 434
QY 141 GlnAlaMetLeuGIuGIuSerThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg 160
Db 435 CAGGCGCATGCTCGGCGACAGACACGAGAGCTGCGGGTGCCTGCCCTCCCACTGGCGC 494
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGIuLysArgLeuAlaValTyr 180
Db 495 AAGCTGCTAAGCGGCTCTCCGCGATGCCGATGACCTGCAGAAAGCGCTGCACTGTAC 554
QY 181 GlnAlaGIuAlaArgGIuGIuValArgGIuArgGIuLeuSerAlaIleArgGIuArgLeuGIu 200
Db 555 CAGGCGGGGCGCGGAGGCGCGGAGGCGGCTGCACGCGCATCCGCGGAGGCGCTGGGG 614
QY 201 ProLeuValGIuGIuGIuArgValArgValArgAlaThrValGIuSerLeuAlaGIuGIuPro 220
Db 615 CCCGTGGTGGAAACAGAGGCGCGTGGGGCGCGCACTGTGGGCTCCCTGGCGGCAAGCGC 674
QY 221 LeuGIuGIuArgAlaGIuAlaThrValArgLeuArgAlaArgMetGIuGIuMetGIu 240
Db 675 CTACAGAGAGCGGCGCGGCGGCGGCGGCGGCTGCGCGCGGATGGAGAGATGGCG 734

QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 735 AGCGGACCGCGACCGCCGTGGACGAGGTGAAGAGAGAGTGGCGGAGGTGGCGCCAG 794
QY 261 LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 277
DB 795 CTGGAGAGAGCGCCAGCATACGCTCGAGCGCCGAGGCTTCCAGGCC 845

RESULT 15
AAF84316
ID AAF84316 standard; cDNA; 1156 BP.
AC AAF84316;
DT 21-JUN-2001 (first entry)
DE Human ApoE4 coding sequence.
KW Human; ApoE4; Alzheimer's disease; arteriosclerosis; ss.
XX Homo sapiens.
FH Key location/Qualifiers
FT CDS 61..1014
FT /tag= a
FT /product= "Human ApoE4"
XX
PN JP2001017028-A.
PD 23-JAN-2001.
XX
PF 28-APR-2000; 2000JP-0128919.
XX
PR 06-MAY-1999; 99JP-0125647.
XX
PA (MITU) MITSUBISHI CHEM CORP.
XX
DR WPI: 2001-285406/30.
DR P-PSDB; AAB80998.
XX
XX New apoe humanized mammalian cell useful for screening for agents
PT useful for treating or preventing Alzheimer's disease and
PT arteriosclerosis -
XX
PS Disclosure; Page 16-17; 22pp; Japanese.
XX
CC The present invention relates to an ApoE humanised mammalian cell. The
CC present sequence is the coding sequence for human ApoE4, which was used
CC in the method of the present invention. The ApoE humanised mammalian cell
CC can be used for screening for agents useful for treating or preventing
CC Alzheimer's disease and arteriosclerosis.
XX
SO Sequence 1156 BP; 208 A; 369 C; 432 G; 147 T; 0 other;

Alignment Scores:
Pred. No.: 2,156-101 Length: 1156
Score: 1362.00 Matches: 276
Percent Similarity: 99.64% Conservative: 0
Best Local Similarity: 99.64% Mismatches: 1
Query Match: 99.13% Indels: 0
DB: 22 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x AAF84316 (1-1156)

QY 1 MetLysValLeuThrPAlaAlaLeuValThrPheLeuAlaGlyGlnAlaLysVal 20
DB 61 ATGAAGGTTCTGTGGGCTGTGTGTCACATTCCTGGCAGATGCCAGGCCAAGGTG 120
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrGluIns 40
DB 121 GAGCAAGCGGTGGACAGAGCGCGAGCCCGAGCTGCCACAGACCGAGTGGCAGAGC 180

QY 41 GlyGlnArgTrpGluLeuAlaLeuGlnGlyArgPheTrpAspTyrIleuArgTrpValGlnThr 60
DB 181 GGCCAGCGCGTGGGAACTGGCAGCTGGGTGCTTTGGATTACCTGGCTGGGTGCAGACA 240
QY 61 LeuSerGluGlnValGlnGluGlnIleuLeuSerSerGlnValThrGlnGlnIleuArgAla 80
DB 241 CTGTCTGAGCAGGTGCGAGAGAGCTGCTCAGCTCCAGGTACCCAGAGAACTGAGAGGCC 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnIleu 100
DB 301 CTGATGGACAGACCATGAAGAGTTGAAGGCTTACAAATCGGAACCTGGAGAACTAG 360
QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
DB 361 ACCCGGTGGCGGAGAGAGAGCGCGGCTGTCCAGAGAGCTGCAGCGGCGCAGGCC 420
QY 121 ArgLeuGlnValAspMetGluAspValLysGlyArgLeuValGlnThrArgGlyGluVal 140
DB 421 CGGCTGGCGCGGACATGGAGAGCTGCGCGGCTGTGTGACATGACGGCGAGGTG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnIleuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCCATGCTGGCCAGAGCACCGAGAGCTCGGGTGGCTCGCTGCCACCTGCCG 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 541 AACCTGCTAAGCGGCTCCCTCCGATGCCGATGACCTCGAGAAAGCGCTGGCAGCTGAC 600
QY 181 GlnAlaGlnValaArgGlnGlnIleuArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 601 CAGGCGGGGCGCGGAGGCGCGGAGCGGCTCAGCGCCATCCGAGCGCGCTGGGG 660
QY 201 ProLeuValGlnGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
DB 661 CCCCTGTGGAAACAGGCGCGCGGCGGCGGCGGCTGCTGGCGCGCGCGCGCGCG 720
QY 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
DB 721 CTACAGAGAGCGGGCCAGGCTGTGGGCGAGCGGCTCGCGCGGAGTGAAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
DB 781 AGCGGAGCCGCGACCGCTGGACGAGGTGAAGAGACAGTGGCGGAGCTGGCGCCAGG 840
QY 261 LeuGlnGlnGlnAlaGlnGlnIleuArgLeuGlnAlaGlnAlaPheGlnAla 277
DB 841 CTGGAGAGAGAGCGCCAGCAGATACGCTGCAAGCGCGAGGCTTCCAGGCC 891

Search completed: March 14, 2003, 12:43:02
Job time : 161.176 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:17:52 ; Search time 30.3514 Seconds
(without alignments)
2798.866 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_277
Perfect score: 1374
Sequence: 1 MKVLMALVTLFLAGCCAKV.....RAKLEGAQOIRLOAEAFQA 277

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPRO.spool/US09827554/rumat_11032003_101611_27510/app_query.fasta_1.3576
-DB-issued=Patents_NA -OPT=fastap -SUFFIX=p2n.rml -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=-1 -MATRIX-blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=Pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827554.gcn.1.1.224 &rumat_11032003_101611_27510 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAR -LARGOQUERY -NEG_SCORES=0 -NAIT -LONGLOG -DEV.TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCrus.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1365	99.3	1157	1	US-07-709-949-1 Sequence 1, Appli
2	1002.5	73.0	1126	4	US-08-949-155-5 Sequence 5, Appli
3	1002.5	73.0	1126	4	US-09-819-964-5 Sequence 5, Appli
4	984	71.6	660	2	US-08-726-306A-28 Sequence 28, Appli
5	866.5	63.1	4267	4	US-08-949-155-51 Sequence 51, Appli
6	866.5	63.1	4267	4	US-09-819-964-51 Sequence 51, Appli
7	511	37.2	330	1	US-07-849-389-6 Sequence 6, Appli
8	374	27.2	252	3	US-08-617-256-24 Sequence 24, Appli
9	374	27.2	252	4	US-09-287-141-24 Sequence 24, Appli
10	374	27.2	252	4	US-09-431-613-24 Sequence 24, Appli
11	374	27.2	252	4	US-09-504-245-24 Sequence 24, Appli
12	374	27.2	252	4	US-09-287-682-24 Sequence 24, Appli

13	374	27.2	252	4	US-09-287-679-24 Sequence 24, Appli
14	374	27.2	252	4	US-09-397-766-24 Sequence 24, Appli
15	374	27.2	252	4	US-09-287-681-24 Sequence 24, Appli
16	374	27.2	252	4	US-09-495-444-24 Sequence 24, Appli
17	161.5	11.8	801	1	US-07-939-946-4 Sequence 4, Appli
18	161.5	11.8	801	1	US-08-333-577-4 Sequence 4, Appli
19	161.5	11.8	801	5	PCT-US92-08634-4 Sequence 4, Appli
20	161.5	11.8	842	4	US-08-952-796-1 Sequence 2, Appli
21	152.5	11.1	5661	4	US-08-938-105-2 Sequence 2, Appli
22	147	10.7	1879	4	US-09-750-550-2 Sequence 2, Appli
23	146.5	10.7	3256	2	US-08-968-751-3 Sequence 2, Appli
24	141.5	10.3	863	1	US-08-448-606-7 Sequence 7, Appli
25	141.5	10.3	964	1	US-08-448-606-5 Sequence 5, Appli
26	141	10.3	6306	1	US-08-195-487-3 Sequence 3, Appli
27	141	10.3	6306	5	PCT-US93-06160-3 Sequence 3, Appli
28	140	10.2	11236	1	US-07-853-913-1 Sequence 1, Appli
29	137.5	10.0	516	4	US-09-183-861-34 Sequence 34, Appli
30	137.5	10.0	516	4	US-09-022-765-34 Sequence 34, Appli
31	137.5	10.0	13121	4	US-08-961-527-126 Sequence 126, App
32	137	10.0	6306	1	US-08-466-390-3 Sequence 3, Appli
33	137	10.0	6306	1	US-08-470-950-3 Sequence 3, Appli
34	137	10.0	6306	1	US-08-467-781-3 Sequence 3, Appli
35	137	10.0	6306	2	US-08-483-924-3 Sequence 3, Appli
36	136	9.9	1771	2	US-08-533-669A-7 Sequence 7, Appli
37	136	9.9	1771	2	US-08-511-872-1 Sequence 1, Appli
38	136	9.9	1771	4	US-09-183-861-7 Sequence 7, Appli
39	136	9.9	1771	4	US-09-022-765-7 Sequence 7, Appli
40	135.5	9.9	4296	4	US-09-060-410-3 Sequence 3, Appli
41	134.5	9.8	603	4	US-08-957-796-14 Sequence 14, Appli
42	132.5	9.6	2085	2	US-08-668-128B-7 Sequence 7, Appli
43	131.5	9.6	2085	2	US-08-905-445-7 Sequence 7, Appli
44	131.5	9.6	8789	1	US-08-328-254-5 Sequence 5, Appli
45	131.5	9.6	10136	1	US-08-353-700-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-07-709-949-1
Sequence 1, Application US/07709949
Patent No. 5472858
GENERAL INFORMATION:
APPLICANT: Attie, Alan D
APPLICANT: Gretch, Daniel G
APPLICANT: Sturley, Stephen L
APPLICANT: Beckage, Nancy E
TITLE OF INVENTION: Production of Recombinant Proteins in
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Quarles & Brady
STREET: P.O. Box 2113
CITY: Madison
STATE: WI
COUNTRY: USA
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/709,949
FILING DATE: 19910604
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 9629691801
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 1:


```

SEQUENCE CHARACTERISTICS:
LENGTH: 1157 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 62..1015
FEATURE:
NAME/KEY: misc.signal
LOCATION: 62..119
PUBLICATION INFORMATION:
AUTHORS: McLean, J W
JOURNAL: Journal of Biological Chemistry
VOLUME: 259
PAGES: 6498-6504
DATE: 1984
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 1157
US-07-709-949-1

Alignment Scores:
Pred. No.: 9,61e-125 Length: 1157
Score: 1365.00 Matches: 275
Percent Similarity: 99.288 Conservative: 0
Best Local Similarity: 99.288 Mismatches: 2
Query Match: 99.348 Indels: 0
DB: 1 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-07-709-949-1 (1-1157)
QY 1 MetLysValLeuTrpAlaIleuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 62 ATGAAGGTTCTGTGGCTGCTGTGCTGCATTCCTGGCAGAGATGCAGGCCAAGGTG 121
QY 21 GUGlnAlaValAGlUthrGluProGluProGluLeuArgGlnGluThrGluTrpGlnSer 40
DB 122 GAGCAAGCGGTGGAGACAGACCGGAGCCGAGCTGCGCAGACAGACCGAGTGGCAGAC 181
QY 41 GylGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB 182 GGCAGCGCTGGGAAGCTGGCTGCTTTGGGATTACTGGCGCGGCGGAGACA 241
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 242 CTGTCTGACACAGGTGCGAGGAGAGCTGCACCTCCCAAGTCACCCAGAACGAGGCGG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGluGlnLeu 100
DB 302 CTGATGACAGACACATGAGAGGAGGCTTCAAAATCGAATCGAGAGAACACATG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaGlnAla 120
DB 363 ACCCGGATGCGGAGAGAGCGGCGGCGCTGTCCAAAGAGCTGAGACGCGCGCAGGCC 421
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
DB 422 CGGCTGGCGCGGACATGAGAGACGTCGCGCGCTGTCAGTACCGCGGCGAGGTG 481
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 483 CAGGCCATGCTGGCCAGAGACACCGAGGAGCTGCGGCTGCGCTCCACCTGCAC 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 542 AAGCTGCGTAAACGGGCTCTCTCCGCGATCCGATGACCTCAGAAAGCGCTGGCAGGTAC 601
QY 181 GlnAlaGlyAlaArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 200

```

```

DB 602 CAGCGCGGGCCCGCGAGGGCCCGGAGCGCGCTCAGGCCATCCGAGGCGCTGGGG 661
QY 201 ProLeuValGluGlnGluValArgValAlaIaIaThrValGlySerLeuAlaGlyGlnPro 220
DB 662 CCCCTGTGTGAACAGGCGCGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
QY 221 LeuGlnGluArgAlaGlnAlaIaIaTrpGlyGluArgLeuArgAlaArgMetGluLysMetGly 240
DB 722 CTACAGAGAGCGGGCCAGGCTGTGGGCGAGCGGCTGCGCGCGATGAGAGATGGGC 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 782 AGTCGAGCCCGGAGACCGCTGAGCAGAGTGAAGAGACAGTGGCGGAGGCTGCCCAAG 841
QY 261 LeuGlnGluGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 277
DB 842 CTGAGAGAGCAGGCCCGACGATACGCTGACAGGCGGAGGCTTCCAGGCC 892

RESULT 2
US-08-949-155-5
Sequence 5, Application US/08949155
Patent No. 6271436
GENERAL INFORMATION:
APPLICANT: Piedrahita, Jorge A
APPLICANT: Bazer, Fuller W
TITLE OF INVENTION: Compositions and Methods for the
TITLE OF INVENTION: Generation of Transgenic Animal Species
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESS: ARNOLD, WHITE AND DURKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/949,155
FILING DATE: Concurrently Herewith
CLASSIFICATION: 800
PRIORITY INFORMATION:
APPLICATION NUMBER: US 60/027,338
FILING DATE: 11-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/046,094
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hibler, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TAWK-177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 51..1001
US-08-949-155-5

Alignment Scores:
Pred. No.: 2.59e-89 Length: 1126
Score: 1002.50 Matches: 198
Percent Similarity: 83.93% Conservative: 37

```


Best Local Similarity: 70.71%
Query Match: 72.96%
DB: 4

Mismatches: 38
Indels: 7
Gaps: 2

US-09-827-854-15_COPY_1_277 (1-277) x US-08-949-155-5 (1-1126)

```

Oy 1 MetLysValLeuThrPalaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 51 ATGAGGGTTCTGTGGCTGTTGGTGTAAACCTCTCCGAGAGATGCCGAGACAGAGAC 110
Oy 21 GluGlnAlaValGluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 111 GAGCGGGG-----CCGCGCGGAGGTGACGCTGTGGTGGAGAGCCCAAG 158
Oy 38 TrpGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrp 57
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 159 TGGCAGGGCAGCAGCCCTGGAGAGCCCTGGGCGCCCTCTGTGGATTACCTGCGCTGG 218
Oy 58 ValGlnThrLeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGlu 77
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 219 GTGCAAGTCCCTGTGTACCAAGTGCAGAGAGAGTGTCTACAGACCAAGTCAACCCAGGAA 278
Oy 78 LeuArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlu 97
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 279 CTGACGAGAGCTGATAGAGAGAGAGCTGAAGAGGAGTGAAGGCTTACCGGAGAGCTGGAG 338
Oy 98 GluGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAla 117
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 339 GCGCAGCTGGGCGCCGTACCAAGAGAGAGAGCGCGCCCTGTCTCAAGAGAGCTGACGCG 398
Oy 118 AlaGlnAlaArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArg 137
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 399 GCGCAGGCGCCGGTGGGGCGGACATGAGAGAGCTGCCAACCCGCTTGTGCTCTACCGC 458
Oy 138 GlyGluValGlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSer 157
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 459 AGCGAGGTGCACACATGTTGGGCGCAGACACCGAGAGAGTGGCGAGCGCTGCTGCC 518
Oy 158 HisLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeu 177
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 519 CACCTGCCCAAGCTGCGAGAGGGCTGCTCCGCGACACCGAGAGAGCTCAAGAGCGCTG 578
Oy 178 AlaValTyrGlnAlaGlyAlaArgGluGluGluAlaGluArgGlyLysSerAlaIleArgGlu 197
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 579 GCGGTGTACAGCGGGGCTGCGGAGGGCGCCGAGCGCAGCGTGTAGGCGCCCTCCGCGAG 638
Oy 198 ArgLeuGlyProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAla 217
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 639 CGCCCTCGGGCGCCCTGTGTGAGCAGCGCCGATGCGCGCCACCTGATACCAAGGCC 698
Oy 218 GlyGlnProLeuGlnGluArgAlaGlnAlaThrProGlyGluArgLeuArgAlaArgMetGlu 237
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 699 GCGCAGCGCTGCGGAGCGGGGAGAGCTGGGGCCAGAGAGCTGCGGAGCGCTGGAG 758
Oy 238 GluMetGlySerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluVal 257
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 759 GAGATGGGACACCGGAGCCCGGAGCCGCTGGATGATGATGCGTGTGAGAGAGAGTGTG 818
Oy 258 ArgAlaLysLeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAla 277
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 819 CGCACCAAAAGTGTGAGAGACGAGCGCAGCTGTGCGCTGTGAGAGCGGATTCACAGCC 878

```

RESULT 3
US-09-819-964-5

Sequence 5, Application US/09819964
Patent No. 6369294

GENERAL INFORMATION:

APPLICANT: Bazer, Fuller W

TITLE OF INVENTION: Compositions and Methods for the

Generation of Transgenic Animal Species

NUMBER OF SEQUENCES: 51

CORRESPONDENCE ADDRESS:

```

ADDRESSER: ARNOLD, WHITE AND DURKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/819,964
FILING DATE: 28-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/949,155
FILING DATE: <Unknown>
APPLICATION NUMBER: US 60/046,094
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hibler, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TAWK:177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 769-2679
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 51..1001
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-819-964-5
Alignment Scores:
Pred. No.: 2.59e-89 Length: 1126
Score: 1002.50 Matches: 198
Percent Similarity: 83.93% Conservative: 37
Best Local Similarity: 70.71% Mismatches: 38
Query Match: 72.96% Indels: 7
DB: 4 Gaps: 2
US-09-827-854-15_COPY_1_277 (1-277) x US-09-819-964-5 (1-1126)
Oy 1 MetLysValLeuThrPalaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 51 ATGAGGGTTCTGTGGCTGTTGGTGTAAACCTCTCCGAGAGATGCCGAGACAGAGAC 110
Oy 21 GluGlnAlaValGluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 111 GAGCGGGG-----CCGCGCGGAGGTGACGCTGTGGTGGAGAGCCCAAG 158
Oy 38 TrpGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrp 57
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 159 TGGCAGGGCAGCAGCCCTGGAGAGCCCTGGGCGCCCTCTGTGGATTACCTGCGCTGG 218
Oy 58 ValGlnThrLeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGlu 77
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 219 GTGCAAGTCCCTGTGTACCAAGTGCAGAGAGAGTGTCTACAGACCAAGTCAACCCAGGAA 278
Oy 78 LeuArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlu 97
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 279 CTGACGAGAGCTGATAGAGAGAGAGCTGAAGAGGAGTGAAGGCTTACCGGAGAGCTGGAG 338
Oy 98 GluGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGlnLeuGlnAla 117
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 339 GCGCAGCTGGGCGCCGTGACCAAGAGAGAGAGCGCGCCCTGTCTCAAGAGAGCTGACGCG 398

```


CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/819,964
FILING DATE: 28-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/949,155
FILING DATE: <Unknown>
APPLICATION NUMBER: US 60/046,094
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hildner, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TAMK:177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 4267 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 51:
US-09-819-964-51

Alignment Scores:
Pred. No.: 3,08e-75 Length: 4267
Score: 866.50 Matches: 192
Percent Similarity: 56.05% Conservative: 35
Best Local Similarity: 47.41% Mismatches: 40
Query Match: 63.06% Indels: 139
Gaps: 4

US-09-827-854-15_COPY_1_277 (1-277) x US-09-819-964-51 (1-4267)

QY 1 MetLysValIleuThrAlaIleuValuThrPheLeuAlaGlySerGlnAlaLysVal 20
Db 2448 ATGGCTGTGAACCTGCTCGCT-----GCAGATGCGGACAGAGAGAC 2489
QY 21 GlnGlnAlaValGlnThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
Db 2490 GAGCGCGGG-----CCGCGCGCGGAGCTGCTGCTGGAGAGAGCCCAAG 2537
QY 38 TrpGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrp 57
Db 2538 TGGCAGGCGACGCGCTGCGAGAGAGCGCCCTGGCGCGCTTGGGATTACCTGGCGCTGG 2597
QY 58 ValGlnThrLeuSerGlnGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlu 77
Db 2598 GTGCACTCTCTGTGACCAAGTGCAGAGAGAGCTGCTCAGCACCAAGTCAACCCAGAA 2657
QY 78 Leu----- 78
Db 2658 CT-GACGTAAGTCCACCCGACTCCCGCGCGCGCGCGCGCGCGCGCGCGCTGA 2716
QY 78 ----- 78
Db 2717 CCTCCTGCGAACCCTGTCTTGACCCCTCAGGCTCCACCGCTCGGGTTTCTTTCTG 2776
QY 78 ----- 78
Db 2777 TCCCTGTGCGCAACTCTTGGGGCTGAGGCTCTGTCTTTCTTTTCTTCTCTCTCTTT 2836
QY 78 ----- 78
Db 2837 TTGGGGGGAATACTTTTCTTTTCTTTTCTTTTCTTTTCACTTACCTGCTTCTTTCC 2896
QY 78 ----- 78
Db 2897 ATCTGAGCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2956
QY 78 ----- 78

Db 2957 AATCTTGCGACGCTCTGCGCATCGCCAGCTCAGAGACCTCTTCTCCCTCACCGCC 3016
QY 79 -----ArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyr 92
Db 3017 CCGGCCCTCTCTGCGCCAGGAGAGTGTATAGAGAGAGATGAAAGAGGCTTGAAGGCTTAC 3076
QY 93 LysSerGluLeuGlnGluGlnLeuThrProValAlaGlnGluThrArgAlaArgLeuSer 112
Db 3077 CCGAGAGACTGTGAGCGGAGGAGCTGGGCGCCGTGACCCAGAGAGAGCGGCGGCTGTC 3136
QY 113 LysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluLysPalaCysGlyArg 132
Db 3137 AAGAGACTCAGAGCGGCGGAGGCGCCCGCTGGCGCGCGGACATGAGAGACTGCGCAACCC 3196
QY 133 LeuValGlnTyrArgGlyGlnValGlnAlaMetLeuGlnSerThrGlnGluLeuArg 152
Db 3197 TTGGTCTCTACCGCAGGAGGTGCACACATGTTGGCCAGACCACGAGAGGTGCGG 3256
QY 153 ValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAsp 172
Db 3257 AGCGGCTGCTGCTCCACCTGCGCAGAGCTGCGCAAGCGGCTGCTCGGACACCGAGAG 3316
QY 173 LeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluValAlaGluArgGlyLeu 192
Db 3317 CTGCAGAACGCGCTGCGCTGTACCAAGCGGCGGCTGCGGAGGCGCGGACGCGTGG 3376
QY 193 SerAlaIleArgGluArgLeuGlyProLeuValGlnGlnGlyArgValArgAlaAlaThr 212
Db 3377 AGCGGCTGCGGAGGAGGCTGCGGCGCGCTGCGGAGGAGGCGGAGGCGCGGCGGCG 3436
QY 213 ValGlySerLeuAlaGlyGlnProLeuGlnGluArgAlaGlnAlaTrpGlyLysArgLeu 232
Db 3437 CTGAGTACCAAGGCGCGGCGGCGGCTGCGGAGCGGCGGAGGCGGAGGCGGAGGAG 3496
QY 233 ArgAlaIleArgMetGlnGluMetGlySerArgTrpArgAspArgLeuAspGluValLysGlu 252
Db 3497 CCGGAGCGCTGAGAGAGTGGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 3556
QY 253 GlnValAlaGluValArgAlaLysLeuGlnGlnAlaGlnAlaGlnAlaIleArgLeuAla 272
Db 3557 CAGCTGAGAGAGGTGGCGACCAAGTGAAGAGAGAGGAGGAGGAGGAGGAGGAGGAG 3616
QY 273 GlnAlaPheGlnAla 277
Db 3617 GAGGAGTTCACGCGC 3631
RESULT 7
US-07-849-389-6
Sequence 6, Application US/07849389
Patent No. 5525493
GENERAL INFORMATION:
APPLICANT: HORNES, Erik
APPLICANT: UHLEN, Mathias
TITLE OF INVENTION: CLONING METHOD AND KIT
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/849,389
FILING DATE: 19920519
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16787/168/DFBC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)836-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 330 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: HUMAN LIPOPROTEIN E GENE
FEATURE:
NAME/KEY: CDS
LOCATION: 1..330
US-07-849-389-6

Alignment Scores:
Pred. No.: 5,71e-42 Length: 330
Score: 511.00 Matches: 107
Percent Similarity: 98.18% Conservative: 1
Best Local Similarity: 97.27% Mismatches: 2
Query Match: 37.19% Indels: 0
DB: 1 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-07-849-389-6 (1-330)

OY 87 LysGluLeuLysAlaTyrLysSerGluLeuGluGluThrProValAlaGluGlu 106
DB 1 AAGGATTGAAGGCTCACAATCGAATCGAGGAGGAACAACCTGCGGCGAGGAG 60
OY 107 ThrArgAlaArgLeuSerLysLysLysGluGlnAlaAlaGlnAlaArgLeuGluAlaSpMet 126
DB 61 ACCGCGGCGACGGCTGTCCAAAGAGCTGCAGCGCGGAGAGCCCGCTGGCGCGCATG 120
OY 127 GluArgValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGluGln 146
DB 121 GAGGAGCTCGCGGCGCGCTGTGCAGTACCGCGGAGAGTGCAGGCCCATGCTCGGCCAG 180
OY 147 SerThrGluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeu 166
DB 181 AACACCGAGAGACTGCGGGGTGGCGCTCCGCTCCCAAGCTGCGTAAGCGGCTC 240
OY 167 LeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGlu 186
DB 241 CTCGCGGATGCCGATGACCTGCAGAACCGCTGCGCAGTGTACCAAGCGCGGGCGCGAG 300
OY 187 GlyAlaGluArgGlyLeuSerAlaIleArg 196
DB 301 GCGCGCGAGCGCGCTCCAGCGCCATCCGC 330

RESULT 8

US-08-617-256-24
Sequence 24, Application US/08617256
Patent No. 6043031

GENERAL INFORMATION:
APPLICANT: Kyster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/617,256
FILING DATE: March 18, 1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: March 17, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth A.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: S01-013CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-617-256-24

Alignment Scores:
Pred. No.: 9.9e-29 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 27.22% Indels: 1
DB: 3 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-08-617-256-24 (1-252)

OY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaArgLeuGlyAlaAspMetGluSp 128
DB 2 GCACGGCTGTCCAAAGAGCTGCAGCGGCGCGAGGCCCGCTGGCGGAGCATGAGAGC 61
OY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGluGlnSerThr 148
DB 62 GTGTGCGC-CGCTGTGTGCAGTACCGCGGAGGTCAGGCGCATCTCGGCCAGAGCAC 120
OY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuArg 168
DB 121 GAGGAGCTCGCGGCTGCGCTCCCTCCCACTGCGCAAGCTGCGTAAGCGGCTCCTCCGC 180
OY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluAla 188
DB 181 GATGCCGATGACCTGCAGAACGCTGCGCAGTGTACCAAGCGCGGGCGCGAGGCGGCC 240
OY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCGCTC 252

RESULT 9

US-09-287-141-24
Sequence 24, Application US/09287141
Patent No. 6197498

GENERAL INFORMATION:
APPLICANT: Kyster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & Maulliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS


```
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,141
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-587-5360
TELEFAX: 619-587-5360
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-141-24

Alignment Scores:
Pred. No.: 9.9e-29 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 27.22% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-287-141-24 (1-252)
QY 109 AlArGleuSerlysgluLeuGlnAlaGlnAlaArgleuGlyAlaAspMetGluAsp 128
Db 2 GCACGGCTGTCCAAAGAGCTGCAGGCGCGCCAGGCCGCTGGGGCGGACATGAGAGAC 61
QY 129 ValCysGlyARgleuValGlnTYrArgGlyGluValGlnAlaMetLeuGlnSerThr 148
Db 62 GTGTGGGC-CGCTGTGGTGCAGTACCGGGGAGGTGCAGGCCATGCTCGGCCAGAGCACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuArg 168
Db 121 GAGGAGCTGGGGTGGCGCCCTCCACCTCCGCAAGCTGGTAAAGGGGCTCTCCGC 180
QY 169 AspAlaAspAspleuGlnLysArgLeuAlaValTYrGlnAlaGlyAlaArgGluGlyAla 188
Db 181 GATGCCGATGACTTCGAGAGAGTCTCTGGCAGTGTACCAAGCGCGGGCCCGGAGGGCGCC 240
QY 189 GluArgGlyLeu 192
Db 241 GAGCGCGGCCTC 252

RESULT 10
US-09-431-613-24
; Sequence 24, Application US/09431613
; Patent No. 6221601
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA

COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/431,613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002G
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-587-5360
TELEFAX: 619-587-5360
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-431-613-24

Alignment Scores:
Pred. No.: 9.9e-29 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 27.22% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-431-613-24 (1-252)
QY 109 AlArGleuSerlysgluLeuGlnAlaGlnAlaArgleuGlyAlaAspMetGluAsp 128
Db 2 GCACGGCTGTCCAAAGAGCTGCAGGCGCGCCAGGCCGCTGGGGCGGACATGAGAGAC 61
QY 129 ValCysGlyARgleuValGlnTYrArgGlyGluValGlnAlaMetLeuGlnSerThr 148
Db 62 GTGTGGGC-CGCTGTGGTGCAGTACCGGGGAGGTGCAGGCCATGCTCGGCCAGAGCACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuArg 168
Db 121 GAGGAGCTGGGGTGGCGCCCTCCACCTCCGCAAGCTGGTAAAGGGGCTCTCCGC 180
QY 169 AspAlaAspAspleuGlnLysArgLeuAlaValTYrGlnAlaGlyAlaArgGluGlyAla 188
Db 181 GATGCCGATGACTTCGAGAGAGTCTCTGGCAGTGTACCAAGCGCGGGCCCGGAGGGCGCC 240
QY 189 GluArgGlyLeu 192
Db 241 GAGCGCGGCCTC 252

RESULT 11
US-09-504-245-24
; Sequence 24, Application US/09504245
; Patent No. 6221605
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
```


NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCulliffe LLP
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/504,245
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002J
TELECOMMUNICATION INFORMATION:
TELEPHONE: 858-450-8400
TELEFAX: 858-587-5360
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-504-245-24

Alignment Scores:
Pred. No.: 9.9e-29
Score: 374.00
Percent Similarity: 97.62%
Best Local Similarity: 97.62%
Query Match: 27.22%
DB: 4
Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-504-245-24 (1-252)

QY 109 AlaArGleuSerLySGluLeuGlnAlaGlnAlaArGleuGlyAlaAspMetGluasp 128
Db 2 GCACGGCTGTCCAGAGACTCAGCGCGCGCGCGCGCGCATGAGAGAC 61

QY 129 ValGysGlyArGleuValGlnTyrArGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
Db 62 GTGTGGCG-CGCTGTGTGTCAGTACCGCGGCGAGGCGCATGCTCGCGCAGACACC 120

QY 149 GluGluLeuArGValArGleuAlaSerHisLeuArGlyLeuArGlyArGLeuLeuArG 168
Db 121 GAGAGAGCTGGCGGTGCGCTGCTCCACCTGCGCAAGCTGCTAAGCGGCTCTCCGC 180

QY 169 AspAlaAspAspLeuGlnLyArGleuAlaValTyrGlnAlaGlyAlaArGLeuGlyAla 188
Db 181 GATGCCATGACCTGCAGAAAGTCCCTGGCACTGTACCAAGCGCGGCGCCGCGAGGCGCC 240

QY 189 GluArGlyLeu 192
Db 241 GAGCGGCGCTC 252

RESULT 12
US-09-287-682-24
; Sequence 24, Application US/09287682
; Patent No. 6235478

GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
APPLICANT: Braun, Andreas
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCulliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-682-24

Alignment Scores:
Pred. No.: 9.9e-29
Score: 374.00
Percent Similarity: 97.62%
Best Local Similarity: 97.62%
Query Match: 27.22%
DB: 4
Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-287-682-24 (1-252)

QY 109 AlaArGleuSerLySGluLeuGlnAlaGlnAlaArGleuGlyAlaAspMetGluasp 128
Db 2 GCACGGCTGTCCAGAGACTCAGCGCGCGCGCGCGCATGAGAGAC 61

QY 129 ValGysGlyArGleuValGlnTyrArGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
Db 62 GTGTGGCG-CGCTGTGTGTCAGTACCGCGGCGAGGCGCATGCTCGCGCAGACACC 120

QY 149 GluGluLeuArGValArGleuAlaSerHisLeuArGlyLeuArGlyArGLeuLeuArG 168
Db 121 GAGAGAGCTGGCGGTGCGCTGCTCCACCTGCGCAAGCTGCTAAGCGGCTCTCCGC 180

QY 169 AspAlaAspAspLeuGlnLyArGleuAlaValTyrGlnAlaGlyAlaArGLeuGlyAla 188
Db 181 GATGCCATGACCTGCAGAAAGTCCCTGGCACTGTACCAAGCGCGGCGCCGCGAGGCGCC 240

QY 189 GIUAArglyLeu 192
| | | | |
Db 241 GAGCGGGCCTC 252

RESULT 13
US-09-287-679-24
; Sequence 24, Application US/09287679
; Patent No. 6258538
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; APPLICANT: Little, Daniel P.
; APPLICANT: Braun, Andreas
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037-9103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/287,679
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 18-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/406,199
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-2002C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-587-5360
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEO ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-287-679-24

Alignment Scores:
Pred. No.: 9.9e-29 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 27.22% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-287-679-24 (1-252)

QY 109 AlaArgLeuSerLyGluLeuGlnAlaAlaArgLeuGlyAlaAspMetGluasp 128
| | | | |
Db 2 GCACGGCTGTCACAGAGCTGCAGCGCGCGCTGCGCGCACATGAGGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
| | | | |
Db 62 GTGTGCGC-CGCTGTGTGATACCGCGCGAGGTGACAGGCATCTCGCCAGAGACACC 120

QY 149 GIUJluLeuArgValArgLeuAlaSerHisLeuArgLyLeuArgLyArgLeuArg 168
| | | | |
Db 121 GAGAGCTCGGGGCTCCCTCCACCTGCCACAGCTCGTAAGCGCTCCCTCC 180

QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
| | | | |
Db 181 GATGCCGATGACCTGCAGAGATCCCTGCGCATGTACAGGCCGGGGCCCGCAGGGCGCC 240

QY 189 GIUAArglyLeu 192
| | | | |
Db 241 GAGCGGGCCTC 252

RESULT 14
US-09-397-766-24
; Sequence 24, Application US/09397766
; Patent No. 6268144
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037-9103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/397,766
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 18-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/406,199
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-2002I
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-587-5360
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEO ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-397-766-24

Alignment Scores:
Pred. No.: 9.9e-29 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 27.22% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-397-766-24 (1-252)

QY 109 AlaArgLeuSerLyGluLeuGlnAlaAlaArgLeuGlyAlaAspMetGluasp 128
| | | | |

DB 2 GCACGGCTGTCCAGAGCTGCAGCGCGCCGCTGGCGCGGACATGAGGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGlnValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTCCGC-CGCTGTGTGACAGTCCGCGGAGGTGCAGGCCATCTCGGCGAGACACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuArg 168
DB 121 GAGGAGCTGCGGGGTGCGCTGCCCTCCACCTGCGGCAAGCTGCTAAGCGGCTCTCCGC 180
QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGTCCCTGGCAGTACAGGCGGCGGCGCGGAGGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGGCGCTC 252

RESULT 15
US-09-287-681-24
Sequence 24, Application US/09287681
Patent No. 6277573
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,681
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-681-24

Alignment Scores:
Pred. No.: 9.9e-29
Score: 374.00
Percent Similarity: 97.62%
Best Local Similarity: 97.62%
Query Match: 27.22%

Length: 252
Matches: 82
Conservative: 0
Mismatch: 2
Indels: 1

DB: 4 Gaps: 0
US-09-827-854-15_COPY_1_277 (1-277) x US-09-287-681-24 (1-252)
QY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
DB 2 GCACGGCTGTCCAGAGCTGCAGGCGCGGCGGCTGGCGGACATGAGGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGlnValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTCCGC-CGCTGTGTGACAGTCCGCGGAGGTGCAGGCCATCTCGGCGAGACACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuArg 168
DB 121 GAGGAGCTGCGGGGTGCGCTGCCCTCCACCTGCGGCAAGCTGCTAAGCGGCTCTCCGC 180
QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGTCCCTGGCAGTACAGGCGGCGGCGGAGGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGGCGCTC 252

Search completed: March 14, 2003, 20:18:31
Job time : 34.3514 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 17:32:58 ; Search time 57.563 Seconds
(without alignments)
3377.460 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_277
Perfect score: 1374
Sequence: 1 MKYIMALLVTFLAGCAK.....RAKLEQAQRIQLQAFAQA 277

Scoring table:
BLOSUM62
Xgapop 10.0, Ygapext 0.5
Xgapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 501302 seqs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09827854/rnat_11032003_101612_27578/app_query.fasta_1.3576
-DB=published.Applications.NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MAPFIX-blosum62
-TRANS=Numan40.cdi -LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100
-THR.MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09827854.ecgn_1.1.298_@rnat_11032003_101612_27578
-NCPU=6 -ICPU=3 -NO_XLIPX -NO_MAP -LARGOQUERY -NEG.SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications.NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Description	
Result No.	Score	Query Match Length	ID		
1	1374	100.0	1156	9 US-09-870-759-129	Sequence 129, App
2	1374	100.0	1156	9 US-09-802-640-17	Sequence 17, Appl
3	1374	100.0	1156	10 US-09-827-854-8	Sequence 8, Appl
4	1374	100.0	1291	12 US-10-044-090-454	Sequence 454, App

5	1370	99.7	1156	10 US-09-827-854-12	Sequence 12, Appl
6	1366	99.4	1156	10 US-09-827-854-9	Sequence 9, Appl
7	1366	99.4	1156	10 US-09-827-854-11	Sequence 11, Appl
8	1365	99.3	1157	10 US-09-954-456-760	Sequence 760, App
9	1365	99.3	1157	10 US-09-880-107-2244	Sequence 2244, App
10	1362	99.1	1156	10 US-09-827-854-7	Sequence 7, Appl
11	1359	98.9	1156	10 US-09-827-854-10	Sequence 10, Appl
12	1158.5	84.3	41907	10 US-09-967-013-5	Sequence 5, Appl
13	656.5	47.8	786	10 US-09-925-302-133	Sequence 133, App
14	625	45.5	478	10 US-09-964-824-374	Sequence 374, App
15	625	45.5	478	10 US-09-880-107-2491	Sequence 2491, App
16	463	33.7	356	10 US-09-960-352-5420	Sequence 5420, App
17	458.5	33.4	449	10 US-09-960-352-4726	Sequence 4726, App
18	439.5	32.0	442	10 US-09-960-352-9395	Sequence 9395, App
19	435	31.7	414	10 US-09-960-352-4237	Sequence 4237, App
20	428.5	31.2	416	10 US-09-960-352-8720	Sequence 8720, App
21	422.5	30.7	423	10 US-09-860-352-14047	Sequence 14047, A
22	405.5	29.5	409	10 US-09-960-352-5148	Sequence 5148, App
23	401.5	29.2	416	10 US-09-960-352-9797	Sequence 9797, App
24	401	29.2	253	10 US-09-179-5368-130	Sequence 130, App
25	387.5	28.2	425	10 US-09-960-352-3497	Sequence 3497, App
26	380.5	27.7	390	10 US-09-960-352-1311	Sequence 1311, App
27	379.5	27.6	391	10 US-09-960-352-1278	Sequence 1278, App
28	379.5	27.6	392	10 US-09-960-352-10599	Sequence 10599, A
29	379.5	27.6	401	10 US-09-960-352-4187	Sequence 4187, App
30	375.5	27.3	413	10 US-09-960-352-8042	Sequence 8042, App
31	374	27.2	252	10 US-09-796-416-24	Sequence 24, Appl
32	374	27.2	252	10 US-09-879-341-24	Sequence 24, Appl
33	369.5	26.9	404	10 US-09-960-352-2325	Sequence 2325, App
34	369.5	26.9	405	10 US-09-960-352-14063	Sequence 14063, A
35	351.5	25.6	377	10 US-09-960-352-176	Sequence 176, App
36	350.5	25.5	353	10 US-09-960-352-4914	Sequence 4914, App
37	350.5	25.5	378	10 US-09-960-352-10040	Sequence 10040, A
38	350.5	25.5	386	10 US-09-860-352-10361	Sequence 10361, A
39	348.5	25.4	388	10 US-09-860-352-11886	Sequence 11886, A
40	344.5	25.1	396	10 US-09-960-352-11886	Sequence 11886, A
41	344.5	25.1	377	10 US-09-960-352-10540	Sequence 10540, A
42	338.5	24.6	432	10 US-09-960-352-8934	Sequence 8934, App
43	338	24.6	332	10 US-09-960-352-6552	Sequence 6552, App
44	337.5	24.6	376	10 US-09-960-352-5408	Sequence 5408, App
45	335.5	24.4	373	10 US-09-960-352-3681	Sequence 3681, App

ALIGNMENTS

RESULT 1
US-09-870-759-129
; Sequence 129, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870, 759
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 129
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(1014)
; OTHER INFORMATION:
US-09-870-759-129

Alignment Scores:
Pred. No.: 1.01e-120
Score: 1374.00
Percent Similarity: 100.00%
Length: 1156
Matches: 277
Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-870-759-129 (1-1156)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 61 ATGAAGGTTCTGTGGCTCGTGTGCTGACATTCTCGCAGAGATGCCAGGCCAAGGTG 120
QY 21 GUGlnAlaValAlaGlnThrLupProGluLeuArgGlnGlnThrLupProGlnSer 40
DB 121 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCCAGACCGAGTGGCGAGCC 180
QY 41 GlyLnaTrpGluLeuAlaLeuGlnArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 181 GGGCAGCGCTGGGAAGTGGCTGGCTTTGGGATTACTGGCTGGGTGGAGACA 240
QY 61 LeuSerGluGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 241 CTGTCTGACAGCTGACAGAGAGCTGCTCAGCTCCAGAGTCAACCGAGACTGAGGGCG 300
QY 81 LeuMetAspGluThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
DB 301 CTGATGACAGACCATGAAGAGTTGAAGCCCTACAAATCGGAATCGAGAGACAACCTG 360
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
DB 361 ACCCGGCTGGCGAGAGACGCGGGCGACGCTGTCCAAAGAGCTGCAGCGCGCGAGCC 420
QY 121 ArgLeuGlnAlaAspMetLysPvalLysGlyArgLeuValGlnTyrArgGlyGlnVal 140
DB 421 CGGCTGGGGCGACATGAGAGACGTGTGCGCGCTGCTCCAGTACCGCGCGAGGTG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCCATCTGCTGGCCAGACACCGAGAGCTGGGGTCCCTCCCTCCACCTGGCC 540
QY 161 LysLeuArgLysArgLeuLeuArgAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 541 AAGCTGCGTAACGGGCTCTCCGCGATGCCGATGACCTCCAGAGCGCTGGCAGGTAC 600
QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 601 CAGCGCGGGCGCGAGGCGCGCGCGCGCTCAGCGCATCCGCGAGCGCGCTGGGG 660
QY 201 ProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
DB 661 CCCCTGGTGAACAGGGCGCGCTGGCGCGCGCTGCTGGCTCCCTGGCGCGCGAGCGG 720
QY 221 LeuGlnGlnArgAlaGlnAlaThrTrpGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
DB 721 CTACAGAGACGGGGCCAGGCTGTGGGCGAGCGGCTGGCGCGATGAGAGAGATGGGC 780
QY 241 SerArgTrpArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
DB 781 ACCCGGACCCCGGACGCGCTGGAGCAGAGTGAAGAGACAGTGGCGAGGTGGCGCGCAG 840
QY 261 LeuGlnGlnGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 277
DB 841 CTGGAGAGACGAGCCAGCATACGCTGTGAGCGCGAGCGCGCTTCAGAGCC 891

RESULT 2

US-09-802-640-17
Sequence 17, Application US/09802640

Publication NO. US20030036057A1

GENERAL INFORMATION:

APPLICANT: Braun, Andreas

APPLICANT: Bonsai Aruna

APPLICANT: Kiehn Patrick

TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND THEIR USE

FILE REFERENCE: 24/36-2048

CURRENT APPLICATION NUMBER: US/09/802,640
CURRENT FILING DATE: 2001-03-09
NUMBER OF SEQ ID NOS: 122
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 17
LENGTH: 1156

TYPE: DNA

ORGANISM: Homo sapien

FEATURE:

NAME/KEY: CDS

LOCATION: (61) ..(1014)

OTHER INFORMATION: Nucleotide Sequence encoding apolipoprotein E

OTHER INFORMATION: (APOE)

US-09-802-640-17

Alignment Scores:

Pred. No.: 1,01e-120

Score: 1374.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 9

DB: 9

Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-802-640-17 (1-1156)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 61 ATGAAGGTTCTGTGGCTCGTGTGCTGACATTCTCGCAGAGATGCCAGGCCAAGGTG 120
QY 21 GUGlnAlaValAlaGlnThrLupProGluLeuArgGlnGlnThrLupProGlnSer 40
DB 121 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCCAGACCGAGTGGCGAGCC 180
QY 41 GlyLnaTrpGluLeuAlaLeuGlnArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 181 GGGCAGCGCTGGGAAGTGGCTGGCTTTGGGATTACTGGCTGGGTGGAGACA 240
QY 61 LeuSerGluGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
DB 241 CTGTCTGACAGCTGACAGAGAGCTGCTCAGCTCCAGAGTCAACCGAGACTGAGGGCG 300
QY 81 LeuMetAspGluThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
DB 301 CTGATGACAGACCATGAAGAGTTGAAGCCCTACAAATCGGAATCGAGAGACAACCTG 360
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
DB 361 ACCCGGCTGGCGAGAGACGCGGGCGACGCTGTCCAAAGAGCTGCAGCGCGCGAGCC 420
QY 121 ArgLeuGlnAlaAspMetLysPvalLysGlyArgLeuValGlnTyrArgGlyGlnVal 140
DB 421 CGGCTGGGGCGACATGAGAGACGTGTGCGCGCTGCTCCAGTACCGCGCGAGGTG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
DB 481 CAGGCCATCTGCTGGCCAGACACCGAGAGCTGGGGTCCCTCCCTCCACCTGGCC 540
QY 161 LysLeuArgLysArgLeuLeuArgAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 541 AAGCTGCGTAACGGGCTCTCCGCGATGCCGATGACCTCCAGAGCGCTGGCAGGTAC 600
QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 601 CAGCGCGGGCGCGAGGCGCGCGCGCGCTCAGCGCATCCGCGAGCGCGCTGGGG 660
QY 201 ProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
DB 661 CCCCTGGTGAACAGGGCGCGCTGGCGCGCGCTGCTGGCTCCCTGGCGCGCGAGCGG 720
QY 221 LeuGlnGlnArgAlaGlnAlaThrTrpGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
DB 721 CTACAGAGACGGGGCCAGGCTGTGGGCGAGCGGCTGCTGGCTCCCTGGCGCGAGGTGGGC 780

QY	Db	QY	Db
241	SeIrrqTrrrArGAsPArGleAsAGlValIyGclucIuValIaGluValArGAlaIyS	261	LeuclugIugIaIaIngIuIleArGleuGInaIaGluAaPheGInaIa
250		277	
781	AGCCGAGACCCCGACCCCTCGAGAGAGTAAAGAGAGAGTGGCGGAGAGTCCGCCCCMAC	841	CTTGAGAGAGCAGGCCACGACGATrAGCGCTTCAGAGCCAGAGGCTTCCAGAGGC
840		891	

RESULT 3

```

: US-09-827-854-8
: Sequence 8, Application US/09827854
: Patent No. US20020123093A1
: GENERAL INFORMATION:
: APPLICANT: Zannits, Vassilis
: APPLICANT: Kyriacos, Kyriakos E.
: TITLE OF INVENTION: Compounds and methods for lowering
: TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia
: FILE REFERENCE: 07180/004003
: CURRENT APPLICATION NUMBER: US/09/827,854
: CURRENT FILING DATE: 2001-04-05
: PRIOR APPLICATION NUMBER: US 09/679,088
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 09/544,386
: PRIOR FILING DATE: 2000-04-06
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 8
: LENGTH: 1156
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-827-854-8

```


; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-9

Alignment Scores:

Pred. No.:	5,69e-120	Length:	1156
Score:	1366.00	Matches:	276
Percent Similarity:	99.64%	Conservative:	0
Best Local Similarity:	99.64%	Mismatches:	1
Query Match:	99.42%	Indels:	0
DB:	10	Gaps:	0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-827-854-9 (1-1156)

Oy 1 MettysValleuTPRALAaleuValThrPheluAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGGTTCTGTGGCGTGGCTGTGGTCACTTCTGGCAGATGCCAGGCCAAGGTG 120
Oy 21 GluGlnAlaValaGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAAGCGGTGAGACAGAGCCGAGCTGCCAGCAGACAGCCGAGTGGCAGAGC 180
Oy 41 GlyGlnArgTTPGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTyrValGlnThr 60
Db 181 GCCCAGCCCTGGAGCACTGGCTGCTGCTTTGGATTAACCTGCTGGGTGCAGACA 240
Oy 61 LeuSerGluGlnValaGluGlnLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGTGTAGCAGAGTCCAGAGAGAGCTGCTCACTCCAGATCCAGAGAACTGAGGGCG 300
Oy 81 LeuMetaspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
Db 301 CTGATGAGCAGAGACCATGAGAGTTGAAGGCTTCAATATGGAACTGGAGAACACTG 360
Oy 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 361 ACCCGGTTGGGAGAGAACGGGGCAGCGCTGTCCAAGAGACTCCAGCGCGCGCAGCGC 420
Oy 121 ArgLeuGlnAlaAspMetGlnAspValaCysGlyArgLeuValaGlnTyrArgGlyVal 140
Db 421 CGCGTGGCGCGGACATGAGAGAGCTGTGGCGCCCTGTGTCACTACCGCGGCGAGGTG 480
Oy 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValaArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTCGGCCAGAGCAGCAGAGAGCTGGCGGTGGCTGCCCTCCACCTGGCG 540
Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTGCTAAGCGGCTCTCCGCGATGCCATGACCTGCAGAAAGTGGCTGGCACTGTAC 600
Oy 181 GlnAlaGlyAlaArgGlnGlyAlaGlnArgGlyLeuSerAlaHisLeuArgGlnLeuGly 200
Db 601 CAGCGCGGGCGCGGAGGGCGCGCAGCGCGCTCAGCGCCATCCGCGAGCGCTGGGG 660
Oy 201 ProLeuValaGluGlnGlyArgValaArgAlaAlaThrValaGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGGTGAAGCAGAGCGCGGTGGCGCGCACTGTGGCTCCCTCGCGCGCGCAAGCG 720
Oy 221 LeuGlnGlnValaArgAlaGlnAlaTyrGlyLysArgLeuArgAlaArgMetGlnGlnMetGly 240
Db 721 CTACAGGAGCGCGCCAGAGCTGTGGCGCAGCGCGCTGCGCGGATGAGAGATGGCG 780
Oy 241 SerArgThrArgAspArgLeuAspGluValaLysGlnGlnValaGlnValaArgAlaLys 260
Db 781 AGCCGAGCCCGGACCGCTGGACAGGTGAAGGAGCGAGGTGGCGAGGTGGCGCCCAAG 840
Oy 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 277

Db 841 CTGAGAGAGAGAGCCAGCCAGATACGCTGCGAGCGCGAGGCTTCCAGAGCC 891

RESULT 7

US-09-827-854-11

; Sequence 11, Application US/09827854
; Patent No. US20020123093A1

GENERAL INFORMATION:

; APPLICANT: Zanis, Vassilis

; TITLE OF INVENTION: Compounds and methods for lowering

; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; PRIOR FILING DATE: 2001-04-05; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06

; SOFTWARE: FastSeq for Windows Version 4.0

; NUMBER OF SEQ ID NOS: 19

; SEQ ID NO 11

; SOFTWARE: FastSeq for Windows Version 4.0

; LENGTH: 1156

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-827-854-11

Alignment Scores:

Pred. No.:	5,69e-120	Length:	1156
Score:	1366.00	Matches:	276
Percent Similarity:	99.64%	Conservative:	0
Best Local Similarity:	99.64%	Mismatches:	1
Query Match:	99.42%	Indels:	0
DB:	10	Gaps:	0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-827-854-11 (1-1156)

Oy 1 MettysValleuTPRALAaleuValThrPheluAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGGTTCTGTGGCGTGGCTGTGGTCACTTCTGGCAGATGCCAGGCCAAGGTG 120
Oy 21 GluGlnAlaValaGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAAGCGGTGAGACAGAGCCGAGCTGCCAGCAGACAGCCGAGTGGCAGAGC 180
Oy 41 GlyGlnArgTTPGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTyrValGlnThr 60
Db 181 GCCCAGCCCTGGAGCACTGGCTGCTGCTTTGGATTAACCTGCTGGGTGCAGACA 240
Oy 61 LeuSerGluGlnValaGluGlnLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGTGTAGCAGAGTCCAGAGAGAGCTGCTCACTCCAGATCCAGAGAACTGAGGGCG 300
Oy 81 LeuMetaspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
Db 301 CTGATGAGCAGAGACCATGAGAGTTGAAGGCTTCAATATGGAACTGGAGAACACTG 360
Oy 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 361 ACCCGGTTGGGAGAGAACGGGGCAGCGCTGTCCAAGAGACTCCAGCGCGCGCAGCGC 420
Oy 121 ArgLeuGlnAlaAspMetGlnAspValaCysGlyArgLeuValaGlnTyrArgGlyVal 140
Db 421 CGCGTGGCGCGGACATGAGAGAGCTGTGGCGCCCTGTGTCACTACCGCGGCGAGGTG 480
Oy 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValaArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTCGGCCAGAGCAGCAGAGAGCTGGCGGTGGCTGCCCTCCACCTGGCG 540
Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTGCTAAGCGGCTCTCCGCGATGCCATGACCTGCAGAAAGTGGCTGGCACTGTAC 600


```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M12529
US-09-880-107-2244

Alignment Scores:
Pred. No.: 7,08e-120 Length: 1157
Score: 1365.00 Matches: 275
Percent Similarity: 99.28% Conservative: 0
Best Local Similarity: 99.28% Mismatches: 2
Query Match: 99.34% Indels: 0
DB: 10 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-880-107-2244 (1-1157)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
   |||||
DB 62 ATGAAGGTTCTGTGGCGCTGCTGTGTCATTCCTGGCAGATGCCAGGCAAGGTG 121

QY 21 GluGlnAlaValAlaGluThrGlnProGluProGluLeuArgGlnGlnThrPrgInser 40
   |||||
DB 122 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCCAGCAGACGACGAGGAGCC 181

QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
   |||||
DB 182 GGCCAGCGCTGGAGACTGGCACTGGTGGATTTGCACTTGGCGGTGGCAGACA 241

QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
   |||||
DB 242 CTGTGTGACAGCTGACAGAGAGCTGCTCAGCTCCCAAGTCACCAAGAACTGAGGGCG 301

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnInleu 100
   |||||
DB 302 CTGATGAGAGAGACCATGAGAGAGTTGAAGGCTTACAAATCGGAAGTGGAGAGAACTG 361

QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
   |||||
DB 362 ACCCGGTAGCGGAGAGAGCGGGCAGCGCTGTCCAAAGACTGCAGAGCGCGCAGCGC 421

QY 121 ArgLeuGlyAlaAspMetClnAspValCysGlyArgLeuValGlnTrpArgGlyGlnVal 140
   |||||
DB 422 CGCGTGGCGGCGAGACATGAGAGAGCTGTCGCGCGCTGTGCTACCGCGGCGAGGTG 481

QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
   |||||
DB 482 CAGGCCATGCTGGCCAGAGACACCAAGAGCTGCGGGTGGCTGCTGCCACTGGCC 541

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
   |||||
DB 542 AAGCTGCGTAAGCGGCTCTCCCGCATCCGATGACCTGCACAAAGCGCTGGCACTGTAC 601

QY 181 GlnAlaGlyAlaArgGlnGlyAlaGlnArgGlyLeuSerAlaIleArgGluArgLeuGly 200
   |||||
DB 602 CAGGCGCGGGCGCGGAGGCGCCAGAGCGCGCTCAGGCGCATCGCGAGGCGCTGGGG 661

QY 201 ProLeuValGlnGlnGlyArgValArgValAlaThrValGlySerLeuAlaGlyLysPro 220
   |||||
DB 662 CCCCTGGTGGAGACAGAGCGCGCTGGCGCGCTGTTGGTCTCTGGCGCGCAGCGG 721

QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyLysArgLeuArgAlaArgMetClnGluMetGly 240
   |||||
DB 722 CTACAGAGAGCGGGCCAGGCGCTGGGCGAGCGGCTGCCCGCGGATGGAGAGATGGGC 781

QY 241 SerArgThrArgAspArgLeuAspClnValLysGlnGlnValAlaGlnValArgAlaLys 260
   |||||
DB 782 ACTCGAGACCGCGACCGCTGTCAGCAGAGGAGAGAGAGGTGGCGAGGTGGCGCCAA 841

QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 277
   |||||
DB 842 CTGGAGAGAGCGCGCCAGAGATGCGCTGAGGCGCGAGGCGCTTCCAGGCGC 892

RESULT 10
```

```
US-09-827-854-7
; Sequence 7, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassilis
; APPLICANT: Kypros, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-7

Alignment Scores:
Pred. No.: 1,35e-119 Length: 1156
Score: 1362.00 Matches: 276
Percent Similarity: 99.64% Conservative: 0
Best Local Similarity: 99.64% Mismatches: 1
Query Match: 99.13% Indels: 0
DB: 10 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-827-854-7 (1-1156)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
   |||||
DB 61 ATGAAGGTTCTGTGGCGCTGCTGTGTCATTCCTGGCAGATGCCAGGCAAGGTG 120

QY 21 GluGlnAlaValAlaGluThrGlnProGluProGluLeuArgGlnGlnThrPrgInser 40
   |||||
DB 121 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCCAGCAGACGACGAGGAGCC 180

QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
   |||||
DB 181 GGCCAGCGCTGGAGACTGGCACTGGTGGATTTGCACTTGGCGGTGGCAGACA 240

QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
   |||||
DB 241 CTGTGTGACAGCTGACAGAGAGCTGCTCAGCTCCCAAGTCCACAGAACTGAGGGCG 300

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnInleu 100
   |||||
DB 301 CTGATGAGAGAGACCATGAGAGAGTTGAAGGCTTACAAATCGGAAGTGGAGAGAACTG 360

QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
   |||||
DB 361 ACCCGGTAGCGGAGAGAGCGGGCAGCGCTGTCCAAAGACTGCAGAGCGCGCAGGCG 420

QY 121 ArgLeuGlyAlaAspMetClnAspValCysGlyArgLeuValGlnTrpArgGlyGlnVal 140
   |||||
DB 421 CGGCTGGCGGCGGACATGAGAGAGCTGCGCGCGCTGTGTCAGTACCGCGGCGAGGTG 480

QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
   |||||
DB 481 CAGGCCATGCTGGCCAGAGACACCAAGAGCTGCGGGTGGCTGCTGCCACTGGCC 540

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
   |||||
DB 541 AAGCTGCGTAAGCGGCTCTCCCGCATCCGATGACCTGCACAAAGCGCTGGCACTGTAC 600

QY 181 GlnAlaGlyAlaArgGlnGlyAlaGlnArgGlyLeuSerAlaIleArgGluArgLeuGly 200
   |||||
DB 601 CAGGCGGGGCGCGGAGGCGCGGAGGCGCGCTCAGCGCGCATCCAGCGCGCTTGGGG 660
```



```
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 661 CCCCTGGTGGACAGGCGCCCGTGGCGGCCACCTGTGGGCTCCCTGGCGCGGACACCG 720
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
DB 721 CTACAGAGAGCGGCGCCAGGCTGGGGGAGAGCGGCTGGCGCGCGGAGATGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaLys 260
DB 781 AGCGGAGCGCGGACCGCCTGGAGAGAGTGAAGGAGCGGCGGAGATGAGATGGC 840
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleArgAla 277
DB 841 CTGAGAGAGAGCGGCCAGCATACGCTGAGCGGCGGAGGCTTCCAGGCC 891

RESULT 11
US-09-827-854-10
; Sequence 10, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassilis
; APPLICANT: Kyriacos, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-10

Alignment Scores:
Pred. No.: 2,59e-119 Length: 1156
Score: 1359.00 Matches: 275
Percent Similarity: 99.288 Conservative: 0
Best Local Similarity: 99.288 Mismatches: 2
Query Match: 98.918 Indels: 0
Gaps: 0
DB: 10

US-09-827-854-15_COPY_1_277 (1-277) x US-09-827-854-10 (1-1156)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 61 ATGAAGTCTCTGGCGTGGCTGTGTCACATTCCTCGAGAGATGCCAGGCGAAGGTG 120
QY 21 GluGlnAlaValAlaGluTrpGluProGluProGluLeuArgGlnIleThrGluTrpGlnSer 40
DB 121 GACCAAGCGGTGAGAGAGCGGAGCGCCGAGCTGGCGCAGAGACCGGAGCGAGAGC 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB 181 GGCACAGCGCTGGGAAGTGGACTGGGTGGCTTTGGGATTACTGGCGTGGGTGGCAGACA 240
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 241 CTGTCTGAGAGAGGTGAGAGGAGAGTGTGCTCAGCTCCAGGTACCCAGAGAACTGAGGGCG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu 100
DB 301 CTGATGAGAGAGCAATGAAGGAGTGAAGGCTTACAAATCGGAATGAGGAGAACTG 360
QY 101 ThrProValAlaGluGluTrpArgAlaArgLeuSerLysGlnGlnAlaGlnAla 120
DB 361 ACCCGGATGGGAGAGAGAGCGGCGACGCTGTCCAAAGAGGCTGAGCGCGCGAGGCC 420
```

```
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
DB 421 CGCGTGGCGCCGAGCATGAGAGACGTGTGGCGGCCCTGTGTGACATCCCGCGAGAGTGC 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGCGCATGCTGACAGAGACCGGAGAGCTCTGGGTGGCTCTGCTCCATCCAGCTGGCG 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 541 AAGCTGGGTAAGCGGCTCTCCCGATGCCGATGACCTGACAGAATGCTCTGGCAGTGTAC 600
QY 181 GlnAlaGlyAlaArgGlnGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 601 CAGGCGGCGCGCGGAGGCGCGGAGCGCGGCTCAGCGCATCCGCGGAGCGCCCTGGGG 660
QY 201 ProLeuValGluGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 661 CCCCTGGTGGACAGGCGCCCGTGGCGGCCACCTGTGGGCTCCCTGGCGCGGACACCG 720
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
DB 721 CTACAGAGAGGCGCCAGGCTGGGGGAGAGCGGCTGGCGCGGAGATGAGATGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaLys 260
DB 781 AGCGGAGCGCGGACCGCCTGGAGAGTGAAGAGAGAGTGGCGGAGTGGCGCGGCGCAAG 840
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleArgAla 277
DB 841 CTGAGAGAGAGCGGCCAGCATACGCTGAGCGGCGGAGGCTTCCAGGCC 891

RESULT 12
US-09-967-013-5
; Sequence 5, Application US/09967013
; Patent No. US20020045840A1
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr, Vincent P.
; APPLICANT: AND USES THEREOF
; TITLE OF INVENTION: METHOD FOR GENETIC ANALYSIS OF APOE DNA
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 11926-022001
; CURRENT APPLICATION NUMBER: US/09/967,013
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: 60/206,613
; PRIOR FILING DATE: 2000-05-23
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 41907
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-967-013-5

Alignment Scores:
Pred. No.: 1.43e-98 Length: 41907
Score: 1158.50 Matches: 260
Percent Similarity: 56.528 Conservative: 0
Best Local Similarity: 56.528 Mismatches: 6
Query Match: 84.328 Indels: 194
DB: 10
Gaps: 1

US-09-827-854-15_COPY_1_277 (1-277) x US-09-967-013-5 (1-41907)
QY 12 PheLeuAlaGlyCysGlnAlaLysValGlnGlnAlaValGlnThrGluProGluProGlu 31
DB 20316 TTCACACAGAGATGCCAGRCCAGGTGAGCAAGCGGTGAGACAGAGCGGAGCGCGAG 20375
QY 32 LeuArgGlnGlnThrGlnTrpGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPhe 51
DB 20376 CTGCGGACAGAGCCAGATGGAGAGCGGCGGAGCGGCTGGGAATGAGTGGCTGCTTT 20435
QY 52 TrpAspTrpLeuArgTrpValGlnThrLeuSerGlnGlnValGlnGlnLeuLeuSer 71
```



```
|||||
Db 20436 TGGGATTACCTGCTGGGTGACAGACTGCTGTGACAGAGTGACAGAGAGCTGCTCAGC 20495
Oy 72 SerGlnValThrGlnLeu-----78
Db 20496 TCCCAAGTACCCAGGAACTGAGTGATGCCATCCGTGACCCTTGACCTCTGCTG 20555
Oy 78 -----78
Db 20556 GCGGCTATACCTCCCAAGTCCAGGTTTCATTCTGCCCTGTGCTAAGTCTTGGGGG 20615
Oy 78 -----78
Db 20616 CCGGGTCTCTGCTGCTTCTAGCTTCTCTTCCCATTTCTGACTCTGGCTTTAGCTCTC 20675
Oy 78 -----78
Db 20676 TGGAAATCTCTCTCTCAGCTTGTCTCTCTCTCTCTCTCTCTCTGACTCAGTCTCACA 20735
Oy 78 -----78
Db 20736 CGTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 20795
Oy 78 -----78
Db 20796 CTCACTGTGTGCCCCAGGCTGTGTTGAACCTTCTGGGCTCAAGCATCTCTCCGCTCGG 20855
Oy 78 -----78
Db 20856 CCTCCCAAGTGTGAGATTAGAGGAGAGACCATCTTGGCCGCTCTGACTCTCTCT 20915
Oy 78 -----78
Db 20916 TCGTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 20975
Oy 78 -----78
Db 20976 GCCTCTGCCCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 21035
Oy 78 -----78
Db 21036 CCCCATCCAGCCCTTCTCCCGCTCCACTGTGCGACACCCCTCCGCTCTCGGCG 21095
Oy 79 -ArgAlaLeuMetAspGlnThrMetTyrGlnValTyrGlnValTyrGlnValTyrGlnVal 21155
Db 21096 CAGGCGCTGATGAGCGAGACCATTAAGAGTGAAGGCTTCAATTTGGAACTGGAGGA 21155
Oy 98 GlnLeuThrProValAlaGlnGluThrArgAlaArgLeuSerTyrGlnValTyrGlnVal 2118
Db 21156 ACAACTGACCCCGGTGGGAGAGAGACCGGCGACGCTGTCCAGAGAGCTGACAGCGGC 21215
Oy 118 aglnAlaArgLeuGlnValAspMetGluAspValCysGlnValTyrGlnValTyrGlnVal 21275
Db 21216 GAGGCGCGGGGGGCGGAGATGAGAGCGTGTGGGCGCGCTGTGCGAGTACCGGCG 21275
Oy 138 GlnValGlnAlaMetLeuGlnValTyrGlnValTyrGlnValTyrGlnValTyrGlnVal 21335
Db 21276 CGAGGTGAGAGCGATGCTCGGCGAGACCGAGAGAGCTGGCGGTGCGCTCGCTCCCA 21335
Oy 158 sLeuArgTyrLeuArgTyrArgLeuLeuArgAspAlaAspSerLeuGlnValTyrGlnVal 21395
Db 21336 CCGTGGCAAGCTGTGTAGCGGCTCTCCGATGCCATGACCTGCGAGAGAGCTCGGC 21395
Oy 178 aValTyrGlnAlaGlnValTyrGlnValTyrGlnValTyrGlnValTyrGlnValTyrGlnVal 21455
Db 21396 AGTGTACCAAGCGCGGGGCGGAGAGCGCGAGCGCGCTCCACCGCATCCGGAGCG 21455
Oy 198 GLeuGlyProLeuValGlnGlnValTyrValTyrAlaTyrValTyrGlnValTyrGlnVal 21515
Db 21456 CCGTGGGCGCGGTGTGAGAGAGCGCGCTGGGCGCGCTGAGAGAGCTGAGAGAGCTGAG 21515
Oy 218 GlnProLeuGlnValTyrGlnValTyrGlnValTyrGlnValTyrGlnValTyrGlnVal 238
```

```
Db 21516 CCAGCCGCTACAGAGAGCGGCGCCAGGCGCTGGGCGAGCGGCTGCGCGCGGATGAGGA 21575
Oy 238 uMetGlySerArgThrArgAspArgLeuAspGlnValTyrGlnValTyrGlnValTyrGlnVal 258
Db 21576 GATGGGAGCGCGGAGCGCGCGAGCGCTGAGAGAGTGAAGAGAGAGTGGCGGAGTGGC 21635
Oy 258 gAlaLysLeuGlnGlnValTyrGlnValTyrGlnValTyrGlnValTyrGlnValTyrGlnVal 277
Db 21636 CGCCAACTGAGAGAGAGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 21693

RESULT 13
US-09-925-302-133
; Sequence 133, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 133
; LENGTH: 786
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-133

Alignment Scores:
Pred. No.: 1,85e-53 Length: 786
Score: 656.50 Matches: 143
Percent Similarity: 97.95% Conservative: 0
Best Local Similarity: 97.95% Mismatches: 3
Query Match: 47.78% Indels: 2
DB: 10 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-925-302-133 (1-786)
Oy 132 ArgLeuValGlnTyrArgGlnValGlnValTyrGlnValTyrGlnValTyrGlnValTyrGlnVal 151
Db 6 CGCGTGTGACAGTACCGGCGGCGAGGTGCASGCGATCTCGGCGAGAGACCGAGAGAGCTG 65
Oy 152 ArgValArgLeuAlaSerHisLeuArgTyrGlnValTyrGlnValTyrGlnValTyrGlnVal 171
Db 66 GGGTGGCGCTCGCTCCACCTGCGCAA-CTGCGTAA-CGGCTCTCTCGGAGATCCGAT 123
Oy 172 AspLeuGlnTyrArgLeuAlaValTyrGlnValTyrGlnValTyrGlnValTyrGlnValTyrGlnVal 191
Db 124 GACCTGCAGAAAGCGCTGCGAGTGTACAGGCGGCGGCGCGAGAGAGAGAGAGAGAGAGAG 183
Oy 192 LeuSerAlaIleArgGlnValTyrGlnValTyrGlnValTyrGlnValTyrGlnValTyrGlnVal 211
Db 184 CTCAGCGCATCCCGGAGAGCGCTGCGGCGCGCTGCGGAGAAAGAGCGCGCTGCGGCGCGC 243
Oy 212 ThrValGlySerLeuAlaGlnValTyrGlnValTyrGlnValTyrGlnValTyrGlnValTyrGlnVal 231
Db 244 ACTGTGGCGCTCCGTGGCGCGCGAGCGCTACAGAGAGCGGCGCGCGCTGCGGCGAGCGCG 303
Oy 232 LeuArgAlaArgMetGlnGluMetGlySerArgThrArgAspArgLeuAspGlnValTyrGlnVal 251
Db 304 CTGCGCGCGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 363
Oy 252 GlnGlnValAlaGlnValTyrGlnValTyrGlnValTyrGlnValTyrGlnValTyrGlnValTyrGlnVal 271
Db 364 GAGCAGGTGGGAGAGGTGCGGCGCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423
Oy 272 AlaGlnAlaIleGlnValTyrGlnValTyrGlnValTyrGlnValTyrGlnValTyrGlnVal 277
Db 424 GCGGAGCGCTTCCAGGCGC 441
```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:09:47 ; Search time 1126.96 Seconds

(without alignments)
3980.771 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_277

Perfect score: 1374

Sequence: 1 MKYLMALLVTFLLGCAQAKV.....RAKLEQAQOIRLQAEAFQA 277

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n model -DEV=xlp

-O=/cgn2_1/USPTO.spool/US09827854/runtat_11032003_101610_27495/app_query.fasta_1.3576

-DB=EST -QMT=fastcat -SUFFIX=p2n.rst -MIMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45

-DOCALLIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000

-USRR=US09827854.cgn.1.1.8826.6runat_11032003_101610_27495 -NCPU=6 -ICPU=3

-NO_XLPXY -NO_MMAR -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120

-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: em_estba:**

2: em_esthum:**

3: em_estlin:**

4: em_estmu:**

5: em_estov:**

6: em_estpl:**

7: em_estro:**

8: em_hnc:**

9: gb_estl:**

10: gb_est2:**

11: gb_hnc:**

12: gb_est3:**

13: gb_est4:**

14: gb_est5:**

15: em_esthum:**

16: em_estom:**

17: gb_gss:**

18: em_gss_hum:**

19: em_gss_inv:**

20: em_gss_pln:**

21: em_gss_vtl:**

22: em_gss_fun:**

23: em_gss_mam:**

24: em_gss_mus:**

25: em_gss_other:**

26: em_gss_pro:**

27: em_gss_rtd:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1254	91.3	842	12	BG763371
2	1225.5	89.2	933	13	B1668318
3	1212.5	88.2	1027	13	B1670367
4	1203	87.6	800	13	BM042094
5	1202.5	87.5	942	13	B1600906
6	1199.5	87.3	922	13	B1597743
7	1190.5	86.6	817	12	BG774871
8	1189	86.5	811	13	B1600553
9	1187.5	86.4	845	12	BG829472
10	1175.5	85.6	927	12	BG472299
11	1171	85.2	790	12	BG707147
12	1152	83.8	919	13	B1551475
13	1149	83.6	757	13	BM042228
14	1139	82.9	907	12	BG706129
15	1129.5	82.2	706	14	BM728696
16	1128	82.1	741	12	BG762924
17	1128	82.1	804	12	BG702752
18	1120.5	81.6	808	13	B1668329
19	1117	81.3	855	13	B1616362
20	1114	81.1	803	13	B1670350
21	1114	80.9	782	12	BG716776
22	1111	80.9	812	13	B1601551
23	1111	80.9	965	14	BQ677266
24	1111	80.9	757	12	B1603658
25	1108	80.5	797	12	BG715366
26	1105.5	80.5	846	13	B1159757
27	1105	80.4	796	13	BM042153
28	1104	80.3	757	12	BG707750
29	1099	80.0	794	13	B1601279
30	1099	80.0	790	13	B1551066
31	1097	79.8	790	13	B1551811
32	1096	79.8	790	12	BG708414
33	1092.5	79.5	798	12	B1549292
34	1091.5	79.4	891	13	B1458355
35	1085.5	79.0	802	13	B1458355
36	1084	78.9	1100	14	BM914382
37	1083.5	78.9	914	13	B1603523
38	1083	78.8	748	13	B1533085
39	1078.5	78.5	812	12	BG769968
40	1077.5	78.4	803	12	BG709360
41	1075	78.2	845	13	B1160936
42	1071	77.9	821	13	B1666579
43	1068.5	77.8	903	13	B1552527
44	1067.5	77.7	810	12	BG714022
45	1067	77.7	757	12	BG703504

ALIGNMENTS

RESULT 1
LOCUS BG763371 842 bp mRNA linear EST 15-MAY-2001
DEFINITION 602735433F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:460585 5',
ACCESSION BG763371
VERSION BG763371.1 GI:14074024
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 842)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

Tissue Procurement: ATCC/DCMD/DFP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

FEATURES

Source

```

1. .842
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image:4860385"
/clone_1id="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/notes="Organ; skin; Vector: pOTR; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCG(g). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC

```

BASE COUNT	154	a	250	c	337	g	101	t
ORIGIN								

ORIGIN

Alignment Scores:

Pred. No.:	4,62e-123	Length:	842
Score:	1254.00	Matches:	260
Percent Similarity:	97.04%	Conservative:	2
Best Local Similarity:	96.30%	Mismatches:	7
Query Match:	91.27%	Indels:	2
DB:	12	Gaps:	0

US-09-827-854-15_COPY_1-277 (1-277) x BG763371 (1-842)

OY	1	MetysValLeuTRPAlaAlaLeuLeuValThrPheLeuAlaGlySGlnAlaVal	20
Db	36	ATGAAGTTCTGTGGGTGCGTTCGTGTCATTCCTGCGAGATGCCAGCCAAAGTG	95
OY	21	GIuGlnAlaValGIuThrGIuProGIuLeuArgGlnInThrGIuTrpInSer	40
Db	96	GAGCAACGCGTGGAGACAGACCAGCCAGCTGGCCAGCAGCAAGTGGCAAGC	155
OY	41	GlyGIuInrArgTRPGIuLeuAlaLeuGlyArgPheTrpAspTryLeuArgTrpValGInThr	60
Db	156	GGCCAGGCGCTGGGAACGTGGCACTGGGTGCTTTGGGATTAACCTGGCCCTGGGTCAACA	215
OY	61	LeuSerGIuGlnValGIuGlnLeuLeuLeuSerSerGlnValThrGlnGIuLeuArgAla	80
Db	216	CTGTCTGAGCAGGTGCAGAGAGACTGCTGCTGACCTCCAGGTACACCGAAGCTAGGGCG	275
OY	81	LeuMetAspGIuTrpMetLeuTrpGIuLeuLysAlaTryIysSerGIuLeuGlnGIuInLeu	100
Db	276	CTGTGTGACAGACCAATGAAGGATTTAAAGCCCTTCAAAATCGGAACGGAGAAACAATG	335
OY	101	ThrProValAlaGIuGIuThrArgAlaArgLeuSerIysGIuLeuGlnAlaAlaGlnAla	120
Db	336	ACCCCGGTGGCGAGAGAGACCGGGGCAAGGCTGTCCAAGACCTGCAGCGGGCGCAGGCC	395
OY	121	ArgLeuGIuLysAspMetGIuAspValCysGIuArgLeuValGlnTryArgGIuVal	140
Db	396	CGCGTGGCGCGGACATGAGAGAGACTGTGGCGCGCTGTGGTCAAGTACCCGCGGAGAGTG	455
OY	141	GlnAlaMetLeuGlnInSerThrGlnGIuLeuArgValArgLeuAlaSerHisLeuArg	160
Db	456	CAGGCATGCTCGGCCAGACACCGAGAGACTGGGTGTGGCGCTTCCTCCACACTTCGCG	515

Oy	181	LysLeuAaGlyVsARLeuLeuAaRgAspAlaAspAlaPleuGlnLbLysARgLeuAaLaValTyr	180
Db	516	AAGCTGCGGTAAAGGGCTCTCTCCGCATGCCATGACCTGCACAAGGGCTGGCAGATGAC	575
Oy	181	GlnAlaGlyAlaAaRgGlnGlyAlaGlnAaRgGlyLeuSerAlaIleAaRgGlnAaRgLeuGly	200
Db	576	CAGCGCGGGGGCCCCGAGAGGGCCCCAGCGCGGCTTCAGCGCCATCCGACAGCGCCCTGGGG	635
Oy	201	ProLeuValGlnGlnGlyARgValARgAlaAlaIleTrnValGlySerLeuAaGlnGlnPro	220
Db	636	CCCTG- GTGGAAACAAGGGCGCGCTGGGGGGCCGAACTGTGGGCTCCCTGGCGGGCAGCGG	694
Oy	221	LeuGlnGlnAaRgAlaGlnAlaIleTrpGlyGlnAaRgLeuAaRgAlaAaRgMetGlnLmet-GI	240
Db	695	CTACGGAGAGCGGGCCACAGGCTGTGGGGCGAGCGGCTGGCGCCCGATGGAGAGATGGCG	754
Oy	240	YSerArGhrArqAspARgLeuAspGlnValLysGlnGlnValAlaGlnValAaIaLys	260
Db	755	CAGCGGAAACCGCGACCGCTGGACGAGGAAAGGAAACCAAGAGTGGGGAGAGTGGCGCAAG	814
Oy	260	SLeuGlnGlnGlnAlaGlnGlnIleAaRg	269
Db	815	CTGGAGAGAGCAGGCGCCAGCAGATATCGC	842

RESULT 2

LOCUS	BI666318	933 bp	mRNA	linear	EST 12-SEP-2007
DEFINITION	603255681f1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:531484 5',				

222

ACCESSION	B1668318
VERSION	PT669319 1

KEYWORDS
VERSION
B10000310.1. 01:155042551
FST

SOURCE

ORGANIC

REFERENCE

AUTHOR

TITLE
TOPIC

COMMENT

CONFIDENTIAL

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shirok
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLML1798 row: a column: 21
 High quality sequence stop: 780.

FEATURES

Sou

```

1. .933
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:531484"
/clone_lid="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
pbluescript="Organ: brain; pbluescript (modified
pbluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gcgcgag
); Oligo-OT primed using primer 5'-ATTTTATTTTATTTTATTTN-3',
size selected for average insert size 2.3 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carroll, In preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH-MGC Library."

```

BASE COUNT	171 a	284 c	359 g
ORIGIN			

ORIGIN

Alignment Scores:		5.74e-120	Length: 933
Pred. No.:	1225.50	Matches: 259	
Score:	94.648	Conservative: 6	
Percent Similarity:	92.508	Mismatches: 12	
Best Local Similarity:	89.198	Indels: 4	
Query Match:	13	Gaps: 1	
US-09-827-854-15_COPY_1_277 (1-277) x B1668318 (1-933)			
QY	1	MethylsValLeuThrPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
DB	75	ATGAAGGTTCTGTGGCGCTGCTGTGTGTCACATTCCTGGCAGATGCCAGGCAAGGTG	134
QY	21	GlUGlnAlaValaGlUthChlupProGluLeuArGlnGlnThChlupPInser	40
DB	135	GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCCAGACACAGAGGCGAGAC	194
QY	41	GIyGlnArGTrPGluLeuAlaLeuGlyArGpHeTrPaSPYrLeuArGTrPaGlnThr	60
DB	195	GGCCAGCGCTGGAGACTGGCACTGGTGGCTTTGGGATTACTGCGTGGGTGCAGACA	254
QY	61	LeuSerGIUGlnValaGlnGluLeuLeuSerSerGlnValThChlUGluLeuArGla	80
DB	255	CTGTCTGAGCAGGTGCAGAGAGAGCTGCTCACCTCCAGGTACCCAGAGAACTGAGGGCG	314
QY	81	LeuMetAspGIuThrMetLysGIuLeuLysAlaTYrLysSerGIuLeuGlnGluGlnLeu	100
DB	315	CTGATGGACGAGACCATGAGAGTTGAGAGCTTACAAATCCGAATCGAGAGACAACTG	374
QY	101	ThrProValAlaGluGluThrArGAlaArGLeuSerLysGIuLeuGlnAlaAlaGlnAla	120
DB	375	ACCCCGGTGGCGAGAGAGCGGGCAGCGCTGTCCAAAGAGACTGACGGCGCGAGGCC	434
QY	121	ArGLeuGIyAlaAspMeGlnAspValaLysGIyArGLeuValaGlnThrArGlyGluVal	140
DB	435	CGGCGGGCGGACATGAGAGAGCTGTGGCGCCCTGTGTGCTACAGCCGGCGAGGTG	494
QY	141	GlnAlaMetLeuGIyGlnSerThrGlnGluLeuArGValaArGLeuAlaSerHisLeuArG	160
DB	495	CAGGCCATGCTGGCCAGAGACCCAGAGAGCTGGGGTGGCTGCTCCCACTGGCC	554
QY	161	LysLeuArGlyArGLeuLeuArGAspAlaAspAspLeuGlnLysArGLeuAlaValTYr	180
DB	555	AAGCTGCGTAAGCGGCTCTCCGCGATGACCTGACAGAACCGCTGGCACTGTAC	614
QY	181	GlnAlaGlyAlaArGLeuGIyAlaGluArGlyLeuSerAlaIlaArGluArGLeuGly	200
DB	615	CAGCGCGGGCGCGGAGGGCGCCAGCGCGGCTCAGCGCCATCCGCGAGCGCTGGGG	674
QY	201	ProLeuValaGluGlnGlyArGValaArGAlaIlaThrValaGlySerLeuAlaGlyGlnPro	220
DB	675	CCCCCTGGTGGAGACAGGGCGGCGGCGGCGGCTGTGGCTCCCTGGCCGAGCGG	734
QY	221	LeuGlnGluArGAlaGlnAlaIlaTrpGIyGluArGLeuArGAlaArG-MeGlu----GluMe	239
DB	735	CTACAGAGGCGGGCCAGGCGCTGGGGCGAGCGGCTGGCGCGCGGATTTGAGACCATTTG	794
QY	239	tGlySerArGThrArGAspArGLeuAsp-GluValLysGIUGlnValaAlaGluValArG	259
DB	795	GGGCGAGCGGAGCCCGCTGAGAGAGGTGAGAGAGAGGTGGGAGTGGCC	854
QY	259	IaLysLeuGluGluGlnAlaGlnGlnIlaArGLeuGlnAlaGluAlaPheGlnAla	277
DB	855	GCCAAAGGCTGGGGGAGACAGGCCCGCAATTGCTTCAGGCCA-CCCTTCACAGGC	909
RESULT 3			
LOCUS	B1670367	1027 bp	mrna linear EST 12-SEP-2001
DEFINITION	603292738F1 NIH_MGC_96	Homo sapiens	cdna clone IMAGE:5312024 5',
ACCESSION	B1670367		
VERSION	B1670367.1	GI:15584600	

KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS	1 (bases 1 to 1027)
TITLE	NIH-MGC http://mgc.ncl.nih.gov/ .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LMML1790 row: 1 column: 09 High quality sequence stop: 845.
FEATURES	Location/Qualifiers
source	1..1027
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="IMAGE:5312024"
	/clone_lib="NIH_MGC_96"
	/tissue_type="hypothalamus"
	/lab_host="DH10B"
	/note="Organ: brain; Vector: pbluescript (modified pbluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcga) ; Oligo-dt primed using primer 5'-TTTTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to 10^5. This is a primary library enriched for full-length clones and constructed using the cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH-MGC Library."
BASE COUNT	194 a 308 c 400 g 125 t
ORIGIN	
Alignment Scores:	
Pred. No.:	1.59e-118 Length: 1027
Score:	1212.50 Matches: 263
Percent Similarity:	93.97% Conservative: 2
Best Local Similarity:	93.26% Mismatches: 10
Query Match:	88.25% Indels: 7
DB:	13 Gaps: 1
US-09-827-854-15_COPY_1_277 (1-277) x B1670367 (1-1027)	
QY	1 MethylsValLeuThrPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB	26 ATGAAGGTTCTGTGGCGCTGCTGTGTGTCACATTCCTGGCAGATGCCAGGCAAGGTG 85
QY	21 GlUGlnAlaValaGlUthChlupProGluLeuArGlnGlnThChlupPInser 40
DB	86 GAGCAAGCGGTGGAGACAGAGCCGAGCTGCCAGACACAGAGGCGAGAC 145
QY	41 GIyGlnArGTrPGluLeuAlaLeuGlyArGpHeTrPaSPYrLeuArGTrPaGlnThr 60
DB	146 GGCCAGCGCTGGAGACTGGCACTGGTGGCTTTGGGATTACTGCGCTGGGTGCAGACA 205
QY	61 LeuSerGIUGlnValaGlnGluLeuLeuSerSerGlnValThChlUGluLeuArGla 80
DB	206 CTGTCTGAGCAGGTGCAGAGAGAGCTGCTCACCTCCAGGTACCCAGAGAACTGAGGGCG 265
QY	81 LeuMetAspGIuThrMetLysGIuLeuLysAlaTYrLysSerGIuLeuGlnGluGlnLeu 100
DB	266 CTGATGGACGAGACCATGAGAGTTGAGAGCTTCAATTCGGAATCGAGAGAACAACTG 325
QY	101 ThrProValAlaGluGluThrArGAlaArGLeuSerLysGIuLeuGlnAlaAlaGlnAla 120

Db 326 ACCCGGTGGCGAGGAGGAGCGGCGACGGCTGTCAAGAGCTGCAGCGCGCAGGCC 385
 QY ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 Db 386 CGGCTGGCGCCGACATGAGGAGACGTGTGGCGCCCTGTGTGACATGCCGCGGAGAGTG 445
 QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 446 CAGGCGCATGCTGGCGCCAGACACCGAGAGAGCTGGGGTGGCGCTCCGCTCCACACTGCCG 505
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 Db 506 AAGCTGGTAAAGCGGCTCCCGCGATGCCATGACCTGCAGAAGCGCCCTGGGACGTGTAC 565
 QY 181 GlnAlaGlyAlaArgGluGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeu-G1 200
 Db 566 CAGGCGCGGCGCCGCGAGGCGCGCGAGCGCGCTCAGCCCATCCCGGAGCGCCCTGGGG 625
 QY 200 YProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPr 220
 Db 626 CCCCCTGTGGAAAGAGTCCGCTGGCGCGCGCCACACTGTGGCTCCCTGGCGCGCAGCC 685
 QY 220 GluGlnGluArgAlaGlnAlaThrPglGluArg-LeuArgAlaArgMetGluGlu-Met 239
 Db 686 GCTACAGAGACGGGCGCAAGGCTGGGGCGAGCGCTTGCAGCGCGATGGAGAGAGC 745
 QY 240 GlySerArg-ThrArgAspArgLeuAspGluValLysGluGlnValAlaGlu-ValArgA 259
 Db 746 GCGAGCGCGGAGCCCGCAGCCCTGACCGACCGAGTGCAAGAGACAGTGGCGGAAAGTCCGC 805
 QY 259 LysLeu---GluGluGlnAlaGlnGln-IleArgLeuGlnAlaGluAlaPhe 275
 Db 806 CCCAAGCTGGACGAGGAGCGCCAGCAAGATACGCTGCGAGCGCGAGGCTTC 859
 RESULT 4
 LOCUS BM042094 800 bp mRNA linear EST 07-NOV-2001
 DEFINITION 603615713F1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5420618 5',
 mRNA sequence.
 ACCESSION BM042094
 VERSION BM042094.1 GI:16771361
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NIH-MGC http://mhc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM1875 row: 1 column: 03
 High quality sequence stop: 792.
 Location/Qualifiers
 1..800
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5420618"
 /clone_lib="NIH_MGC_112"
 /tissue_type="melanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTB7; Site: 1; XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAC(G). Library constructed by Ling Hong in the

BASE COUNT 146 a 239 c 315 g 100 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,13e-117 Length: 800
 Score: 1203.00 Matches: 244
 Percent Similarity: 99.20% Conservative: 4
 Best Local Similarity: 97.60% Mismatches: 2
 Query Match: 87.55% Indels: 1
 DB: 13 Gaps: 0
 US-09-827-854-15_copy_1_277 (1-277) x BM042094 (1-800)
 QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 50 ATGAAAGTTCTGTGGGCTCGGTGTGTACATTCCTGGAGAGTGCAGGCAAGGCTG 109
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGlnPglIns 40
 Db 110 GAGCAAGCGGTGAGACAGACGCGGAGCCGAGCTGCGCACACAGACGAGTGGCAGAC 169
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 Db 170 GCGCAGCGCTGGGAACGTGGACACTGGGTGCTTTGGGATTACTGGCGGTGGAGACA 229
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 Db 230 CTGTCTGACAGAGTGCAGAGAGAGCTGTACACTCCACAGTCAACCAGAACTGAGGGCG 289
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 Db 290 CTGATGAGAGACCACTGAAGAGTTGAAGGCTTACAAATCCGAATCGAAGAACACTG 349
 QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 Db 350 ACCCGGCTGGCGAGAGAGACGCGGCGACGCTGTCCAAAGACACTGACGCGCGCAGGCC 409
 QY 121 ArgLeuGlyAlaAspMetLysAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 Db 410 CGGCTGGCGCGACATGAGAGACGTGTGCGCGCTGTGTCAGTACCGCGCGCAGAGTG 469
 QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 470 CAGGCGCATGCTGGCGCCAGAGACCGAGAGAGCTGGGGTGGCTGCCCTCCACCTGGCC 529
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 Db 530 AAGCTCGTAAAGCGGCTCTCCGCGATGCGATGACCTGCAGAAAGCGCTGCGCAGTGC 589
 QY 181 GlnAlaGlyAlaArgGluGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db 590 CAGGCGCGGCGCGCGAGGCGCCGAGCGCGCTCAGCGCCATCCGCGAGCGCTGGGG 649
 QY 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPr 220
 Db 650 CCCCTGTGTGAACAGAGCGCGCTGCGGCGCACACTGTGGCTCCCTGGCGCGGCAAGCCG 709
 QY 221 LeuGlnGluArgAlaGlnAlaThrPglGluArgLeuArgAlaArgMetGluGlnMetGly 240
 Db 710 CTACAGAGACGCGGCGCAGGCTGTGGGCGAGCGGTG-CGCGCGCGATGAGAGAGATGGGC 768
 QY 241 SerArgThrArgAspArgLeuAspGluVal 250
 Db 769 ACGCGGAGCGGGAACGCTGTGAGCAGCTG 798
 RESULT 5
 LOCUS B1600906 942 bp mRNA linear EST 07-SEP-2001
 DEFINITION 603249241F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5301010 5',

ACCESSION	BRNA sequence.	
VERSION	BI600906	
KEYWORDS	BI600906.1 GI:15493845	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS	1 (bases 1 to 942)	
TITLE	NH-MGC http://mgc.ncl.nih.gov/.	
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN) DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) cDNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://Image.llnl.gov Plate: LLAM11762 row: a column: 11 High quality sequence stop: 762.	
FEATURES	Location/Qualifiers	
SOURCE	1..942 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:J5301010" /clone_lib="NH MGC 96" /tissue_type="hypothalamus" /lab_host="DH10B" /note="Organ: brain; Vector: pBluescript (modified pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgac); Oligo-dt primed using primer 5'-tttttttttttttttVN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NH_MGC Library."	
BASE COUNT	177 a 284 c 361 g 120 t	
ORIGIN		
Alignment Scores:		
Pred. No.:	1.62e-117	Length: 942
Score:	1202.50	Matches: 261
Percent Similarity:	94.35%	Conservative: 6
Best Local Similarity:	92.23%	Mismatches: 10
Query Match:	87.52%	Indels: 7
DB:	13	Gaps: 1
US-09-827-854-15_COPY_1_277 (1-277) x BI600906 (1-942)		
QY	1 MetlysalLeuTPpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	73 ATGAAGGCTTCGTGGGCGTGGCTGTCACATTCCTGGCAGATGCCAGGCCAAGGTG	132
QY	21 GlnGlnAlaValAlaGluThrGluProGluLysProGluLeuAlaGlnGlnThrgluTrrGlnSer	40
Db	133 GAGCAAGCGGTGGAGACAGACCCGAGCCGCGAGCTGCCACCAACGAGGCGAGAC	192
QY	41 GlyGlnATGTPGluLeuAlaLeuGluLysTrpPheTrpAspTyrLeuArgTrpValGlnThr	60
Db	193 GGGCAGCCCTGGGAACCTGGCACCTGGGTGCTTTTGGATTACCTCGCTGGTGGCAGACA	252
QY	61 LeuSerGlnGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80
Db	253 CTGTCTGACAGTGCACAGAGAGAGCTGCTCAAGCTCCACAGTCCACCAAGAACTGAGGCG	312
QY	81 LeuMetAspGluThrMetLysGluLeuLysAlaTrrLysSerGlnLeuGlnGlnGlnLeu	100
Db	313 CTGATGGACGAGACCATTAAGAGATTGAAGGCTTCAATCGGAATCGGAATGAGACAACTG	372

Accession	Source	Organism	Reference	Author	Title	Journal	Comment
101	ThrpProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaGlnAla	120					
373	ACCCTGGTGGCGGAGAGAGCGCGGACCGCTGTCTCCAGAGACTCCAGCGCGCAGGCC	432					
121	ArgLeuGlnAlaAspMetCysGlyArgLeuValGlnTyrArgGlyGluVal	140					
433	CGGCTGGGGCGGACATGGAGAGCGTGTGGCGCGCTGTGCAGTACCGCGGCGAGGTG	492					
141	GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg	160					
493	CAGGCGATGCTCGCGCAGAGCACCCAGAGAGCTGCGGCTGCGCTGCCACCTCGCGC	552					
161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180					
553	AAGCTGCGTAAACGGGCTCTCTCGCGCATGCGGATGACCTGCAGAACCGCTGGCAGGTAC	612					
181	GlnAlaGlnAlaArgGluGluValaGlnArgGlyLeuSerAlaIleArgGluArgLeuGly	200					
613	CAGCGCGGGGCGCGGAGAGGCGCGCAGCGCGGCTTCAGCGGCATCCGCGAGCGCTGGGG	672					
201	ProLeuValGluGlnGlyArgValAlaArgAlaAlaThrValGlySerLeuAlaGlyLysPro	220					
673	CCCTGGTGAACAGGCGCGCTGGCGGCG -GCCACTGTGGCTCTCCCTGGCGCGCAGCGC	731					
221	LeuGlnGluArgAlaGlnAlaThrArgGlyGluArgValaArgAla -ArgMetGlu -GluMetG	240					
732	CTACAGGAGAGCGGGCCAGGCTGTGGGGCCAGCGGGGTGGCGCGCGCGGATGGAGCGATGG	791					
240	LysSerArgThrArgAspArgLeuAsp -GluValLys -GluGlnValAlaGluVal -ArgA	259					
792	GCAGCGCGAGCCCGGACGCGCTGTGACCGAGCGGAGAACGACAGGTGGGGAAGTTGGCGG	851					
259	LalysLeuGlnGluGlnAla---GlnGlnIleArgLeuGlnAlaGluAlaPheGlnAla	277					
852	CCAACTGTGAGAGAACACGCGCCAGCAGATACGCTTGGCAGGCGCGAGCGCTTCAGGCC	910					
RESULT 6							
BI597743							
LOCUS	BI597743	922 bp	mRNA	linear	EST 07-SEP-2001		
DEFINITION	603248609p1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5300259 5'						
ACCESSION	BI597743						
VERSION	BI597743.1	GI:15490682					
KEYWORDS	EST.						
SOURCE	human.						
ORGANISM	Homo sapiens						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.						
TITLE	NIH-MGC http://mgc.nci.nih.gov/.						
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)						
COMMENT	Unpublished (1999)						
	Contact: Robert Strausberg, Ph.D.						
	Email: cgabbs@email.nih.gov						
	Tissue Procurement: Miklos Palkovits, M.D., Ph.D.						
	CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki						
	Toshitsuki and Piero Carninci (RIKEN)						
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)						
	DNA Sequencing by: Incyte Genomics, Inc.						
	Clone distribution: MGC clone distribution information can be						
	found through the I.M.A.G.E. Consortium/LNLN at:						
	http://image.lnl.gov						
	plate: LAM11760 row: b column: 04						
	High quality sequence stop: 782.						
FEATURES	location/Qualifiers						
source	1..922						
	/organism="Homo sapiens"						
	/db_xref="taxon:9606"						
	/clone="IMAGE:5300259"						
	/clone_id="NIH_MGC_96"						
	/tissue_type="hypothalamus"						
	/lab_host="DH10B"						


```

Oy 61 LeuSerGIuGIuValGIuGIuLeuLeuSerGIuValThrGIuLeuAraGla 80
    |||||
Db 220 CTGCTGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 279
Oy 81 LeuMetAspGIuThrMetLysGIuLeuLysAlaTYrLysSerGIuLeuGIuGIuLeu 100
    |||||
Db 280 CTGATGAGCAGAGACCATGATGAGAGGCTTCAAAATCGAAGTGGAGGAGCAACTG 339
Oy 101 ThrProValAlaGIuGIuThrArgAlaArgLeuSerLysGIuLeuAlaGIuAla 120
    |||||
Db 340 ACCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 399
Oy 121 ArgLeuGIuAlaAspMetGIuAspValCysGIuArgLeuValGIuTYrArgGIuVal 140
    |||||
Db 400 CGCCTGGGCGCGGACATGGAGAGCTGTGGCGCGCCCTGTGCATACCGCGGAGAGTG 459
Oy 141 GluAlaMetLeuGIuGIuSerThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg 160
    |||||
Db 460 CAGGCGCATGCTGGCCAGAGCACCGAGAGAGCTGGCGCTGCCCTCCACCTGGCGC 519
Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGIuLysArgLeuAlaValTYr 180
    |||||
Db 520 AAGCTGCGTAAGCGGCTCTCCGCGATGCGCATGACCTGCAGAAAGCGCTGGAGTGTAC 579
Oy 181 GluAlaGIuAlaArgGIuGIuValGIuArgGIuLeuSerAlaIleArgGIuArgLeuGIu 200
    |||||
Db 580 CAGGCGCGGCGCGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 639
Oy 201 ProLeuValGIuGIuGIuArgValArgAlaAlaThrValGIuSerLeuAlaGIuGIuPro 220
    |||||
Db 640 CCCCTGTGTGAACAGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 698
Oy 221 LeuGIuGIuArgAlaGIuAla--TTPGIuGIuArgLeuArgAlaArgMetGIuGIuMetG 240
    |||||
Db 699 CTACAGAGGCGGCGCGCGCGCTTGTGGGCGAGCGGATGGCGCGGATGGAGGAGTGC 758
Oy 240 LysArgThrArgAspArgLeuAspGIuValLysGIuGIuValGIuAlaGIuValArgAla 259
    |||||
Db 759 GGCACCGGACCG--GAACCGCTGACAGAGTGAAGAGCAGAGTGGCGGAGGTGCCGCGC 815

```

```

RESULT 8
LOCUS      B1600563              811 bp      mRNA      linear      EST 07-SEP-2001
DEFINITION 603244936P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5287329 5',
            mRNA sequence.
ACCESSION  B1600563
VERSION    B1600563.1 GI:15493502
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE  1 (bases 1 to 811)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cga@pds-remail.nih.gov
            Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1A11726 row: 9 column: 10
            High quality sequence stop: 783.

```

```

FEATURES
SOURCE     1..811
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5287329"

```

```

/clone_id="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site:1: BamHI; Site:2: SalI-XhoI (gcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to R0.5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH-MGC Library."

```

```

BASE COUNT  148 a 245 c 316 g 102 t
ORIGIN

```

Alignment Scores:

```

Pred. No.: 3,566-116      Length: 811
Score: 1189.00           Matches: 241
Percent Similarity: 98.37%
Best Local Similarity: 97.97%      Mismatches: 1
Query Match: 86.54%           Indels: 1
DB: 13                      Gaps: 0

```

```

US-09-827-854-15_COPY_1_277 (21-277) x B1600563 (1-811)

```

```

Oy 1 MetLysValLeuThrPAlaAlaLeuValThrPLeuAlaGIuCysGIuAlaLysVal 20
    |||||
Db 74 ATGAAGCTTCTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 133
Oy 21 GluAlaValAlaGIuThrGIuProGIuProGIuLeuArgGIuGIuThrGIuTrpGIuSer 40
    |||||
Db 134 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTGGCGCAGACAGACGAGGAGAGC 193
Oy 41 GlyIuAlaGTTTPIuLeuAlaLeuGIuArgPIuPAspTYrLysArgTYrValGIuThr 60
    |||||
Db 194 GGCACGCGCTGGGAACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 253
Oy 61 LeuSerGIuGIuValGIuGIuLeuLeuSerSerGIuValThrGIuGIuLeuArgAla 80
    |||||
Db 254 CTGTCTGAGCAGAGTGCAGAGAGAGCTGCTACCTGCCAGGTACCCAGGACTGAGAGGCG 313
Oy 81 LeuMetAspGIuThrMetLysGIuLeuLysAlaTYrLysSerGIuLeuGIuGIuLeu 100
    |||||
Db 314 CTGATGAGCAGAGACCATGATGAAGGCTTCAAAATCGAAGTGGAGGAGCAACTG 373
Oy 101 ThrProValAlaGIuGIuThrArgAlaArgLeuSerLysGIuLeuGIuAlaGIuAla 120
    |||||
Db 374 ACCCGGTGGCGGAGAGACGCGGCGCGCTGTCCAAGAGACTGCAGGCGCGGAGGCGC 433
Oy 121 ArgLeuGIuAlaAspMetGIuAspValCysGIuArgLeuValGIuTYrArgGIuVal 140
    |||||
Db 434 CGGCTGGGCGCGGACATGAGAGAGCTGTGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGT 493
Oy 141 GluAlaMetLeuGIuGIuSerThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg 160
    |||||
Db 494 CAGGCGCATGCTGGCCAGAGCACCGAGAGCTGGCGGTGGCGCTGCCCTCCACCTGGCGC 553
Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGIuLysArgLeuAlaValTYr 180
    |||||
Db 554 AAGCTGCGTAAGCGGCTCTCCGCGATGCGGATGCTGCACAAACGCGCTGGCAGTGTAC 613
Oy 181 GluAlaGIuAlaArgGIuGIuValGIuArgGIuLeuSerAlaIleArgGIuArgLeuGIu 200
    |||||
Db 614 CAGGCGCGGCGCGCGGAGGCGCGGAGCGCGGCTCAGAGGCACTCCGAGGCGCTGGGG 673
Oy 200 yProLeuValGIuGIuGIuArgValArgAlaAlaThrValGIuSerLeuAlaGIuLys 220
    |||||
Db 674 CCCCTTGTGTGAACAGGCGCGGCTGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 733
Oy 220 GluGIuGIuArgAlaGIuAlaThrPIuGIuArgLeuArgAlaArgMetGIuGIuMetGI 240
    |||||
Db 734 GCTACAGAGAGCGGCGCGGCGCGGCTGGCGGAGGCTGCGGCGGAGTGAAGCAATGGG 793

```


QY 240 ySerArgThrArgAsp 245
|||||
DB 794 CAGCCGGAACCGCGAA 809

RESULT 9
BG761746
LOCUS BG761746 938 bp mRNA linear EST 15-MAY-2001
DEFINITION 602717942F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:484141 5',
mRNA sequence.
ACCESSION BG761746
VERSION BG761746.1 GI:14072399
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE NIH-MGC http://mhc.ncl.nih.gov/
AUTHORS 1 (bases 1 to 938)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CML674 row: c column: 12
High quality sequence stop: 767.
Location/Qualifiers
1. 938
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:484141"
/clone_lib="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-CDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

BASE COUNT 171 a 272 c 374 g 121 t
ORIGIN

Alignment Scores:
Pred. No.: 6,33e-116 Length: 938
Score: 1187.50 Matches: 265
Percent Similarity: 93.388 Conservative: 9
Best Local Similarity: 92.338 Mismatches: 3
Query Match: 86.438 Indels: 10
DB: 12 Gaps: 3

US-09-827-854-15_COPY_1_277 (1-277) x BG761746 (1-938)

QY 1 MetLysValLeuThrAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
|||||
DB 57 ATCAAGGTTCTGTGGGCTGTGTCACATCTCTGGCAGAGATGCCAGCGCAAGG 116

QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
|||||
DB 117 GAGCAAGCGGTGGAGACAGACGCCGCGACCTCGCCAGACGACGAGTGCGCAGAC 176

QY 41 GlyGlnArgTrpGluLeuAlaLeuGluArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
DB 177 GGCACAGCGCTGGGAAGTGGGCTGGTGGCTTTGGGATTACTCGCGCTGGGTCAGACA 236

QY 61 LeuSerGluGlnValGlnGluGluLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
DB 237 CTGTGTGAGCAGAGGTGCAGGAGAGCTGCTCAGCTCCAGTCCAGCAAGCAAGAGCGCG 296

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
|||||
DB 297 CTATGTGACAGACACCATGAGAGGAGTTGAAAGCCCTACAAATCGGAATGGAGCAACACTG 356

QY 101 ThrProValAlaGluGluThrThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
|||||
DB 357 ACCCGGTGGCGGAGAGACAGCGGGCAGCGCTGTCCAAAGAGCTGCAGCGCGGCAAGGCC 416

QY 121 ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
|||||
DB 417 CGGCTGGCGCGGACATGAGACGTGTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 476

QY 141 GlnAlaMetLeuGluGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
DB 477 CAGGCGATGCTCGGCCACAGACACCGAGAGCTGCGGGTGGCTGCTGCTGCTGCTGCTGCTG 536

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
DB 537 AAGCTCGTAAGGCGGCTCTCCGCGATGCCGATGACCTGCAGAGAGCGCTGGCAGTGTAC 596

QY 181 GlnAlaGlyAlaArgGluGluGluArgGlyLeuSerAlaLeuArgGluArgLeuGly 200
|||||
DB 597 CAGCGCGGGCGCGCAGAGGCGCGCGCGCTGCTGAGCGCCATCCGCGCAGCGCGCTGGGG 656

QY 201 ProLeuValGluGlnGluArgValArgAlaAlaThrValGlySerLeuAlaGlyGln 220
|||||
DB 657 CCCCTGTGGACAGGCGCGCGCTGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 716

QY 220 OleuGlnGluArgAlaGlnAlaThrPheGlyGluArgLeuArgAlaArgMetGlu 238
|||||
DB 717 GCTACAGAGAGCGGCG 776

QY 239 MetGlySerArgThrArgAspArgLeuAspGluValLysGluGlnValAla 256
|||||
DB 777 TGGGCG 836

QY 257 ValArgAlaLysLeuGluGluGlnAlaGlnGlnLeuArgLeuGlnAla 275
|||||
DB 837 TTGGCGCGCAAGCTGGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 896

QY 275 eGlnAla 277
|||||
DB 897 CCAGCGCC 903

RESULT 10
BG829472 845 bp mRNA linear EST 22-MAY-2001
LOCUS BG829472
DEFINITION 602763768F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4899112 5',
mRNA sequence.
ACCESSION BG829472
VERSION BG829472.1 GI:14177059
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE NIH-MGC http://mhc.ncl.nih.gov/
AUTHORS 1 (bases 1 to 845)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: L1CM1790 row: g column: 17
High quality sequence stop: 829.

FEATURES

source

Location/Qualifiers

1..845

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4899112"

/clone_lib="NIH_MGC_42"

/tissue_type="epithelioid carcinoma cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT

153 a 244 c 342 g 106 t

ORIGIN

Alignment Scores:

Pred. No.: 1.03e-114

Score: 1175.50

Percent Similarity: 96.23%

Best Local Similarity: 95.47%

Query Match: 85.55%

DB: 12

Length: 845

Matches: 253

Conservative: 2

Mismatches: 7

Indels: 5

Gaps: 2

US-09-827-854-15_COPY_1_277 (1-277) x BG829472 (1-845)

QY 1 MettysValleuThrpAlaAlaLeuValThrpheleuAlaGlyCysGlnAlaLysVal 20
 DB 57 ATGAAGTTCTGTGGCTGCGTTCGTCGATTCCTGGCAGATGCCAGGCAAGGTG 116
 QY 21 GlnGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 117 GAGCAAGCGGTGGAGACAGAGCGCGGAGCGCGGAGCGGAGCGGAGCGGAGCG 176
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpPaspTrpLeuArgTrpValGlnThr 60
 DB 177 GGCACGCGGTGGAGACGTCGCTGCTTGGGATTAACCTCGCGGTGCAGACA 236
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB 237 CTGTGTGACGAGTGCAGAGAGCTGCTCACTCCAGGTACCCAGAGACTGAGGGCG 296
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGluLeu 100
 DB 297 CTGATGGACGACCATGATGAGAGTTGAAGCGCTCAAAATCGAAGTGGAGGACACTG 356
 QY 101 ThrProValAlaGluGluThrArgAlaArgLysSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 357 ACCCGGTTGGCGAGAGAGCGGCGGCGGCTGTCCAAGGAGACTCGAGCGCGCGAGCGCC 416
 QY 121 ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
 DB 417 CGGCTGGCGCGGACATGAGGACGTGTGGCGCGCTGGTGTCACTACCGCGGCAAGGTG 476
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 477 CAGGCGATGCTGCGCGAGAGCACCGAGAGTGGGGGTGGCGCTGCCACTGGCGC 536
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValLys 180
 DB 537 AAGCTCGCTAACCGGCTCTCCGCGATGCGGATGACCTGCAGAAACGCGCTGCGACTGTAC 596
 QY 181 GlnAlaGlyAlaArgGluGlyValArgLysArgLysSerAlaLysArgLysLeuGly 200
 DB 597 CAGGCGCGGGCGCGCGAGGCGCGCGAGCGCGCTTCAGCGCCATCCGAGAGCGCTGGGG 656
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyLysPro 220

DB 657 CCCCCTGGTGAACAGAGCGCGCGTGGCGCGG-ACGTGGGGCTCCCTTGGCGGAG-CCG 714

QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyLysArgLeuArgAlaArgMetGluGluMetGly 240

DB 715 CTACAGGAGCGCGCGAGCGCTG---GGGCGAGCGGCGCGCGCGGATGAGAGATGGGC 771

QY 241 SerArgTrp-ArgAspArgLeuAspGluValLysGluGlnGlnValAlaGlnValArgAlaLys 260

DB 772 AGCCGAGCCGCCA---CGCTGGACGAGGTGAAGACAGATGGCGGAGGTGCGGCCCA 828

QY 260 sLeuGlnGluGln 264

DB 829 GCTTGAGAGCAG 841

RESULT 11

BG472299

LOCUS

602513830F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4645759 5',

DEFINITION

mRNA sequence.

ACCESSION

BG472299

KEYWORDS

EST.

SOURCE

ORGANISM

human.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: image.lnl.gov
 Plate: L1CM1419 row: k column: 08
 High quality sequence stop: 848.

FEATURES

source

Location/Qualifiers

1..927

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4645759"

/clone_lib="NIH_MGC_16"

/tissue_type="retinoblastoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Library constructed by Ling Hong

in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT

170 a 271 c 371 g 115 t

ORIGIN

Alignment Scores:

Pred. No.: 3.53e-114

Score: 1171.00

Percent Similarity: 93.93%

Best Local Similarity: 93.57%

Query Match: 85.23%

DB: 12

Length: 927

Matches: 262

Conservative: 1

Mismatches: 10

Indels: 9

Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x BG472299 (1-927)

QY 1 MettysValleuThrpAlaAlaLeuValThrpheleuAlaGlyCysGlnAlaLysVal 20
 DB 47 ATGAAGTTCTGTGGCTGCGTTCGTCGATTCCTGGCAGATGCCAGGCAAGGTG 106


```
QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluThrGlnSer 40
    |||
Db 107 GAGCAACGGGTGGAGACAGACCGGACCCGAGCTGGCCAGACGAGCCGATGGCAGAGC 166
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
    |||
Db 167 GGGCAGCGCTGGAGACTGGACTGGCTCTTTGGGATTTACTGCTCCCTGGGTGCAGACA 226
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
    |||
Db 227 CTGTCTGAGCAGAGTGCAGAGGAGAGCTCTCAGCTCCAGATCCAGGAACTGAGGGCG 286
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluLeu 100
    |||
Db 287 CTGATGACACAGACCAATGAGAGAGTTGAGAGCCCTACAAATCGGAACTGGAGAACTG 346
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
    |||
Db 347 ACCCGGGTGGCGGAGAGACGGGGCAGCTCTCCAAAGAGCTGCAGGGCGCCAGGCC 406
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
    |||
Db 407 CGGCTGGCGCCGAGACATGAGAGAGCTGTGGCGCCCTGGTGCAGTACCGCGGAGAGTG 466
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
    |||
Db 467 CAGGCGATCTCGGCGCCAGACACCGAGAGCTGGGGCTCGCCCTCCACCTCGCG 526
QY 161 LysLeuArgLysArgLeuArgAlaAspAlaAspLeuGlnLysArgLeuAlaValTyr 180
    |||
Db 527 AAGCTGGTAAGAGGGCTCTCCCGAGTCCGATGACCTGCAGAAAGCCCTGGCAGTGTAC 586
QY 181 Gln-AlaGlyAlaArgGluGluGlyAlaGluArgGly--LeuSerAlaIleArgGluArgLeu 199
    |||
Db 587 CAGGGCGGGGGCGCGGAGGGCGCGAGCGGCCCTCAAGCCGACCGCGAGCGGCTT 646
QY 200 --GlyProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGly 219
    |||
Db 647 GGGGCCCTCTGGTGGAAACAGGGCGCGCTGGCGCCGCTGAGTGGGCTCCCTGGCGCGCC 706
QY 219 LnpProLeuGlnGluArgAlaGlnAlaThrPglLysGluArgLeuArgAlaArgMetGlnGlu 239
    |||
Db 707 AGCG-CTACAGAGAGCGGGCCAGCGCTGGGGCCAGACGGTG-CGCGCGCGGATGAGAGAGA 764
QY 239 etGlySerArg-ThrArgAspArgLeuAspGluValLysGluGlnValAlaGlu-ValAr 258
    |||
Db 765 TGGGGCAGCGGAGACCGCGAGACCTGGACCGAGGTGAAGAGACAGAGTGGCGGAGGTGCG 824
QY 258 gAlaLysLeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlu 273
    |||
Db 825 CCGCAAGCTGGAGAAAGCAGCGCAGAGATACGCTGCAGGCGGAG 870
RESULT 12
Bg707147 790 bp mRNA linear EST 07-MAY-2001
LOCUS Bg707147
DEFINITION 602670283F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4793084 5',
mRNA sequence.
ACCESSION Bg707147
VERSION Bg707147.1 GI:13983201
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 790)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
COMMENT Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Miklos Palcovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
```

```
FEATURES
    source
        1..790
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="IMAGE:4793084"
            /clone_lib="NIH_MGC_96"
            /tissue_type="hypothalamus"
            /lab_host="DH10B"
            /note="Organ: brain; Vector: pBluescript (modified
            pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
            ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
            size-selected for average insert size 2.3 kb and
            normalized to ROT 5. This is a primary library enriched
            for full-length clones and constructed using the
            Cap-trapper method (Carninci, in preparation). Library
            constructed by M. Brownstein (NHGRI/NHGR), National
            Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 143 a 234 c 311 g 102 t
ORIGIN
Alignment Scores:
Pred. No.: 2,94e-112 Length: 790
Score: 1152.00 Matches: 236
Percent Similarity: 98.74% Conservative: 0
Best Local Similarity: 98.74% Mismatches: 3
Query Match: 83.84% Indels: 1
DB: 12 Gaps: 0
US-09-827-854-15_COPY_1_277 (1-277) x Bg707147 (1-790)
QY 1 MetLysValLeuThrPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||
Db 73 ATGAGGTTCTGTGGGTGGTGTGTGTGTCATTCCTCGGACAGATGCCAGGCCAGAGTGC 132
QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluThrGlnSer 40
    |||
Db 133 GAGCAACGGGTGGAGACAGACCGGAGCCGACCTCGCCAGAGACCGAGTGGCAGAGC 192
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
    |||
Db 193 GGGCAGCGCTGGAGACTGGACTGGCTCTTTGGGATTTACTGCTCCCTGGGTGCAGACA 252
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
    |||
Db 253 CTGTCTGAGCAGAGTGCAGAGGAGAGCTGTGCTCAGCTCCAGATCCAGGAACTGAGGGCG 312
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluLeu 100
    |||
Db 313 CTGATGACACAGACCAATGAGAGAGTTGAGAGCCCTACAAATCGGAACTGAGAACTG 372
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
    |||
Db 373 ACCCGGGTGGCGGAGAGACCGCGGACCGCTCTCCAGAGAGTGCAGGGCGGAGGCC 432
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
    |||
Db 433 CGGCTGGCGCCGAGACATGAGAGAGCTGTGGCGCCCTGGTGCAGTACCGCGGAGAGTG 492
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
    |||
Db 493 CAGGCAATGCTGGCCAGAGACCGAGAGGCTCGGGGTCCGCTCCCACTGGCCGCGC 552
QY 161 LysLeuArgLysArgLeuArgAlaAspAlaAspLeuGlnLysArgLeuAlaValTyr 180
    |||
Db 553 AAGCTCGTAAAGCGGCTCTCCCGATGCCGATGACCTGCAGAAAGCGCTGGCAGTGTAC 612
```


QY 181 GlnAlaGlyAlaArgLuglAlaArgLysLeuSerAlaIleArgLuglArgLeuGly 200
|||||
Db 613 CAGCGCGGGCCGCGAGGCGCGGCTCAGCGCCATCGCGAGGCGCTGGGG 672
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
|||||
Db 673 CCCCTGGTGGAAACAGGCGCGCTGGCGCGCCACTGTGGCTCCCTGGCCAGCGAGCCG 732
QY 221 LeuGlnGluArgAlaGlnAlaThrpgLysArgLeuArgAlaArgMetGluGlnMet 239
|||||
Db 733 CTACAGGACCGGCGCCAGGCTGGGCGGAGCGGTG-CGCGCGGATGGAGGATG 788
RESULT 13
LOCUS B1551475 919 bp mRNA linear EST 05-SEP-2001
DEFINITION 603194314F1 NIH_MGC_95 Homo sapiens CDNA clone IMAGE:5274003 5',
mRNA sequence.
ACCESSION B1551475
VERSION B1551475.1 GI:15438787
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS 1 (bases 1 to 919)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ULNL at:
http://image.llnl.gov
Plate: LLM11691 row: 1 column: 04
High quality sequence stop: 812.
FEATURES
source
location/Qualifiers
1..919
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5274003"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size selected for average insert size 2.5 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 172 a 270 c 363 g 114 t
ORIGIN
Alignment Scores:
Pred. No.: 7.61e-112 Length: 919
Score: 1149.00 Matches: 247
Percent Similarity: 93.63% Conservative: 3
Best Local Similarity: 92.51% Mismatches: 16
Query Match: 83.62% Indels: 4
DB: 13 Gaps: 0
US-09-827-854-15_COPY_1_277 (1-277) x B1551475 (1-919)
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||

Db 73 ATGAAGTTCTGTGGCGTGGCTTGCTGTGTCACATTCTGGCAGAGTCCAGGCAAGGTG 132
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGluThrGluProGluInser 40
|||||
Db 133 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCGCCACACAGCCAGTGGCAGAGC 192
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlnArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
Db 193 GGCACAGCGGTGGAACTGACATGCGTGGCTTTTGGATTTACCTGCGGTGGTGCACA 252
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
Db 253 CTGTGTGACAGCTGACAGAGAGAGCTGCTCACCTCCAGAGTCCAGAGAACTGAGGGCG 312
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnGlnLeu 100
|||||
Db 313 CTGATGGACAGACCATGAGAGAGTTGAAGGCTTCAAAATCGAACTGGAGAACACTGG 372
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
|||||
Db 373 ACCCGGTGGCGGAGAGAGAGCGGGGACGCGCTGTCCAGGAGACTGCAGGCGGCGAGCGC 432
QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
|||||
Db 433 CGGCTGGGCGCGACATGAGAGAGAGCTGTGCGGCGCCCTGTGCTACCTACCGCGCAGGCTG 492
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 493 CAGGCGCATGCTGGCCAGAGACCGAGAGAGCTGGCGGTGGCTGCTCCCTCCCACTGGCGC 552
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
Db 553 AAGCTGGCGTAACGGGCTCTCCGCGATGCCATGACATCGCAAAACGCGCTGGCAGTATGAC 612
QY 181 GlnAlaGlyAlaArgGlnGlyAlaGluArgLysLeuSerAlaIleArgLuglArgLeuGly 200
|||||
Db 613 CAGGCGGGGGCGCGATGGCGCCAGCGCGTGTGACAGGCCATCGCGAGCGCGCTGGGG 672
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
|||||
Db 673 CCCCTGGTGGAAACAGGCGCGCTGGCGCGCCACTGTGGCTCCCTGGCGCGAGCCCT 732
QY 221 LeuGlnGluArgAlaGlnAlaThrpgLysArgLeuArgAlaArgMetGluGlnMet 240
|||||
Db 733 A--CAGAGAGCGGGCGCAGGCGCTGGGGCGAGCGGTGCGCGGATGGAGAAATGGG 790
QY 240 ySerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
|||||
Db 791 CAGCGGAGCCCGGAGCG--CTGGACGAAGTGAAGAGACAGCTGGCGGAGGG-TGGCGCAA 847
QY 260 sLeuGlnGluGlnAlaGln 266
|||||
Db 848 GCTGGAGGACCGACGAGAG 866
RESULT 14
LOCUS BM042228 757 bp mRNA linear EST 07-NOV-2001
DEFINITION 603616186F1 NIH_MGC_112 Homo sapiens CDNA clone IMAGE:5557004 5',
mRNA sequence.
ACCESSION BM042228
VERSION BM042228.1 GI:16771495
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS 1 (bases 1 to 757)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: DCTD/DTP


```
OY 21 GIUGInAlaValGluThrGluProGluProGluLeuArgGluGlnThrGluTrpGlnSer 40
    |||||||
Db 135 GAGCAAGCGGTGGACAGACAGCGGAGCCGAGCTGCCACAGACAGACGAGGCGACAGC 194
    |||||||
OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
    |||||||
Db 195 GGCACAGCGGTGGAACTGCTGCTTTGGGATTACCTGCGCTGGGTGCAGACA 254
    |||||||
OY 61 LeuSerGluGlnValGluGluLeuLeuSerSerGlnValThrGluGluLeuArgAla 80
    |||||||
Db 255 CTGTCTGAGCAGGTGACAGAGAGGCTGCTCAGCTCCAGGTCACCCAGAACTGAGGGCG 314
    |||||||
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
    |||||||
Db 315 CTGATGAGACAGACCATGATGAGAGTTGAAGCGCTTACAAATCGAACTGAGAGACACTG 374
    |||||||
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGlu-LeuGlnAlaAlaGlnAl 120
    |||||||
Db 375 ACCCGGTGGCGGAGAGACGCGGCGCTGCTCAAGGAAGCTGCAGGCGGCGCAGGC 434
    |||||||
OY 120 ArgGluGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
    |||||||
Db 435 CCGGCTGGGCGGAGATGAGAGAGTGTGGCGGCCCTGGGTGACGTACCGCGCGAGGT 494
    |||||||
OY 140 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
    |||||||
Db 495 GCAGGCCATGCTCGGCCAGACAGACGAGAGAGTGGGGTGGCCCTCGCTCCACCTGCG 554
    |||||||
OY 160 GlyLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
    |||||||
Db 555 CAAGCTGCGTAAGCGGCTCTCGCGGATGCCGATGACCTGCAGAAAGCGCTGGCAGTGTA 614
    |||||||
OY 180 rGlnAlaGlyAlaArgGlyGluGlyAlaGlyLysLeuSerAlaIleArgGluArgLeu-G 200
    |||||||
Db 615 CCAGGCCGCGGCGCCGAGGGCGGCGAGCGGCGCTCAGCGCCATCCCGAGCGGCTGG 674
    |||||||
OY 200 LysProLeuValGluGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlyIle 220
    |||||||
Db 675 GCCCTTGGTGGACAGAGGCGCGCTGCCGCGCCGACCTGTGGGCTCCCTGNGCGGCAGC 734
    |||||||
OY 220 rOLeuGlnGluArgAlaGlnAlaTrpGlyLysArgLeuArgAlaArgMetGluGluMetG 240
    |||||||
Db 735 CGCTACAGAGAGCGGCGGCGTGG---GGCGCAGCGGTGCGCGCGGATGGAGAGATGG 791
    |||||||
OY 240 LysArgTyrThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArg 258
    |||||||
Db 792 GCAG-CGGACCGCG---ACGCTGAGCAGAGTG-AAAGAGACAGTGGGGAAGTGTGCCG 842
    |||||||
```

Search completed: March 14, 2003, 20:13:55
Job time : 1129.96 secs

